

The
University
Of
Sheffield.

You don't say!

Lying, asserting and insincerity

By:

Neri Marsili

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

The University of Sheffield
Faculty of Arts and Humanities
Department of Philosophy

July 2017

NB: parts of the thesis have been omitted because they are (or will be) under review. For a full version of the thesis, please contact me at n.marsili@shef.ac.uk

Abstract

This thesis addresses philosophical problems concerning improper assertions. The first part considers the issue of defining lying: here, against a standard view, I argue that a lie need not intend to deceive the hearer. I define lying as an insincere assertion, and then resort to speech act theory to develop a detailed account of what an assertion is, and what can make it insincere.

Even a sincere assertion, however, can be improper (e.g., it can be false, or unwarranted): in the second part of the thesis, I consider these kinds of impropriety. An influential hypothesis maintains that proper assertions must meet a precise epistemic standard, and several philosophers have tried to identify this standard. After reviewing some difficulties for this approach, I provide an innovative solution to some known puzzles concerning this issue. In my view, assertions purport to aim at truth, but they are not subject to a norm that requires speakers to assert a proposition only if it is true.

Table of Contents

| | |
|---|-----------|
| <i>I. An introduction</i> | 11 |
| 1. Lies | 12 |
| 2. Incorrect Assertions | 16 |
| 3. Outline of the Dissertation | 19 |
| | |
| <i>Part A- Lying</i> | 21 |
| <i>A. The definition of lying</i> | 23 |
| 1. The Classic Definition of Lying | 23 |
| 2. The Current Debate: Deceptionist vs Non-Deceptionist accounts | 31 |
| <i>II. Deceptionist definitions of lying: a critique</i> | 35 |
| 1. The intention to deceive condition: an introduction | 35 |
| 2. Different versions of the intention to deceive condition | 36 |
| 2.1. The “broad” IDC | 36 |
| 2.2. The “rigid” IDC | 37 |
| 2.3. The “believed sincerity condition” and the disjunctive account | 38 |
| 3. Against sufficiency: non-assertoric falsehoods that intend to deceive..... | 42 |
| 4. Against necessity: lying without the intent to deceive | 46 |
| 4.1 Lying under duress | 46 |
| 4.2 Bald-faced lies! | 47 |
| 5. Deceptionist Replies..... | 50 |
| 5.1. Are bald-faced lies intended to deceive? | 50 |
| 5.2 Are bald-faced lies assertions? | 53 |
| 6. Conclusions..... | 59 |
| <i>III. Lying: a speech-act theoretic approach</i> | 61 |
| 1. Assertion-based definitions, and their difficulties | 62 |
| 1.1. Carson | 63 |
| 1.2. Fallis..... | 65 |
| 1.3. Stokke..... | 68 |
| 2. Three Puzzles From Speech Act Theory | 71 |
| 2.1 Explicit Performatives | 72 |
| 2.2. Determining which speech acts can be lies | 74 |
| 2.3. Insincerity Conditions..... | 80 |
| 3. A Speech-act Theoretic Account of Lying..... | 82 |
| 3.1 The Assertion Condition | 82 |
| 3.2 Illocutionary Entailment | 85 |
| 4. Testing the definition | 87 |
| 4.1 Proviso Lies..... | 87 |
| 4.2 The Speech-act Theoretic Puzzles | 89 |
| 5. Conclusions..... | 94 |

| | |
|---|------------|
| IV. Insincerity | 97 |
| 1. Insincerity: a preliminary account | 99 |
| 2. Beyond belief: insincerity and illocutionary acts | 101 |
| 2.1 The third speech act theoretic puzzle..... | 101 |
| 2.2 Expressing attitudes and (in)sincerity..... | 102 |
| 2.3 Insincerity conditions for promising..... | 104 |
| 2.4 A general account of the insincerity conditions for lying | 108 |
| 3. Insincere Promises: an experimental study | 110 |
| 3.1 Testing folk intuitions about lying..... | 110 |
| 3.2 Aim of the study | 111 |
| 3.3 The predictions of existing theories..... | 112 |
| 3.4 Experiment 1 | 113 |
| 3.5 Experiment 2 | 117 |
| 3.6 Experiment 3 | 118 |
| 3.7 General discussion..... | 120 |
| 3.8. More on the results | 120 |
| 3.9 Conclusions | 125 |
| 4. Gradedness | 127 |
| 4.1 The dichotomic view and the traditional insincerity condition | 127 |
| 4.2 Graded Truth Values | 129 |
| 4.3 Graded beliefs..... | 132 |
| 4.4 A “graded” definition of insincerity | 133 |
| 4.5 Expressing graded beliefs and graded truth values..... | 138 |
| 4.6 Further graded components..... | 144 |
| 5. Conclusions – A general account of insincerity | 148 |
| | |
| Part B - The norms of assertion | 153 |
| B. The norms of assertion | 155 |
| V. A ‘constitutive’ norm? | 161 |
| 1. The orthodox account of constitutive norms | 161 |
| 2. Williamson on constitutive norms | 163 |
| 3. The C-Rule as a Constitutive Rule | 165 |
| 3.1. Phrasing the C-rule as orthodoxly constitutive | 165 |
| 3.2. An alternative constitutive reading | 166 |
| 3.3 Treating two rules as one | 167 |
| 4. The C-Rule as a Regulative Rule | 169 |
| 4.1 The C-rule as orthodoxly regulative..... | 169 |
| 4.2 Pollock’s paradigm: ‘prescriptive’ constitutive rules..... | 169 |
| 4.3 All norms are constitutive in Pollock’s sense | 170 |
| 4.4 The distinction is not meaningful | 174 |
| 5. Conclusions | 175 |
| VI. Truth and assertion: rules vs aims | 179 |
| 1. The disagreement about the norm of assertion | 179 |
| 1.1 The norm of assertion..... | 179 |
| 1.2 Intuitions and norms | 181 |
| 1.3 Falsity-Criticism and inadvertent violations | 182 |
| 2. Lucky and unlucky assertions | 184 |
| 2.1 Lucky: inadvertent observation of TR | 184 |
| 2.2 Unlucky: inadvertent violation of TR/KR | 185 |

| | |
|--|-------------------|
| 2.3 Checkmating the speaker: a too demanding norm..... | 187 |
| 3. Deriving permissibility from correctness | 189 |
| 4. Truth as aim..... | 192 |
| 4.1 The difference between rules and aims | 193 |
| 4.2. Truth as a rule, truth as an aim | 194 |
| 4.3 The truth-aim account | 197 |
| 4.4 Ought to try - from aims back to rules | 199 |
| 5. Objections, replies and clarifications | 202 |
| 5.1 The source of normativity | 202 |
| 5.2 Primary and secondary violations..... | 203 |
| 5.3 Challenges | 208 |
| 5.4 Knowledge as the aim? ² | 208 |
| 5.5 Alleged asymmetries | 209 |
| 6. Conclusions..... | 211 |
| <i>VII. Conclusions.....</i> | <i>213</i> |
| <i>References</i> | <i>231</i> |

Ad Angiolino, Basilio e Riccardo

I. An introduction

It will be convenient to start from the beginning. *In principio erat Verbum, et Verbum erat apud Deum, et Deus erat Verbum* (Jhn, 1, 1-3): “In the beginning was the Word, and the Word was with God, and the Word was God”. If in the beginning of time language was one with God and belonged to God, who is the source of all Truth (Pr, 8,7; Sam 7, 28; Rm 3, 4), its evil relative, the lie, appeared short after creation.

According to the Biblical myth, God told Adam: “Of every tree of the garden thou mayest freely eat: But of the tree of the knowledge of good and evil, thou shalt not eat of it: for in the day that thou eatest thereof thou shalt surely die” (Gen, 2, 16-17). But Satan, “the father of lies”, who “speaks his native language [when he lies]” (Jhn, 8, 44), under the fake appearance of a snake, told Eve: “Ye shall not surely die: For God doth know that in the day ye eat thereof, then your eyes shall be opened, and ye shall be as gods, knowing good and evil” (Gen 3, 4-6). If we believe the Bible, this was the very first lie ever told; and it was this lie that determined the expulsion of humans from paradise, and the beginning of our history.

In the Bible, truth and truthfulness associated with God and the Good; falsity and untruthfulness with the Devil and Evil. The precept “thou shalt not bear false witness against thy neighbour” (do not lie) figures in the Tenth Commandments that God gives to Moses on mount Sinai. The Bible is one of the foundational books of Western culture: it is not a coincidence that truthful and untruthful communications are central themes in this text, given their importance for humanity, and many are the parables and allegories that involve them. In a less allegoric fashion, and with more modest intentions, this dissertation will present and analyse, from a philosophical point of view, the opposition between these two opposite communicative forces.

Communication is a fundamental ability of human beings – it is one of the abilities that makes us distinctively human. Unfortunately, the ability to communicate comes at a price: it makes it possible (and easy) for speakers to misrepresent what they believe, to communicate something false; in other words, it allows communicators to lie. This concept is beautifully illustrated in this famous quote from Umberto Eco (1975):

[The study of communication] is in principle the [study of] everything which can be used in order to lie. If something cannot be used to tell a lie, conversely it cannot be used to tell the truth: it cannot in fact be used "to tell" at all.

Any signal that can be used to tell the truth can be used to tell a falsehood: the ability to communicate sincerely is essentially entangled with the ability to communicate insincerely. To some, it may appear that language is for this reason an intrinsically unreliable source of knowledge. After all, lying is not the only way in which language can deceive us. Not only speakers can be insincere: they can also be mistaken about what they say, and communicate falsehoods even when they are speaking in good faith. And still, despite the possibility of lies and mistakes, we believe most of what other people say.

Rather than a matter of mere gullibility, the trust we place in each other's reports is the result of our dependency on human communication for survival. Most of the knowledge we use to lead our daily lives comes from what others tell us. It is by trusting teachers and authors of books that we apprehended everything that we learnt in school: from physics to geography, from grammar to the whole history of the world. Even our own names are something we learn from hearsay, as none of us can recollect those initial days of our life in which we were given one. Without the trust we place in communication, we would be paralyzed as epistemic agents, and it would be almost impossible for us to lead a meaningful life. And still, this does not mean that other people's testimony is always reliable: it simply means that we cannot afford but to trust most of what we are told.

Considering these issues, it becomes quite clear that understanding truthful and untruthful communication is of central importance, both for the study of language and that of our ordinary life. The aim of this dissertation is to explore untruthful communication from a philosophical point of view. The dissertation is divided in two parts; with some approximation, we could say that each part touches the two sides of this coin: the first part (A) deals with the deliberate communication of something false, the second part (B) with the accidental communication of something false. In what follows, I will clarify more in detail what this means, and to what extent this description is an approximation.

1. Lies

The first part of this dissertation (Part A) will cover philosophical problems related to lying, insincerity and deception. These phenomena are of fundamental importance in contemporary society, where communication plays an increasingly important role. To see

this, it will be helpful to look at some recent examples. If we limit the attention to politics, the current political debate has seen an unprecedented rise in lying and other forms of deceptive communication. Before the Brexit referendum in June 2016, UK media have infamously reported numerous false or deceptive claims. Among these, the false assertion that the NHS was “nearly at breaking point” due to “a massive influx of EU immigrants”, and that “more than 700 offences are being committed by EU immigrants every week”¹. The most blatant of these lies was the pledge to convert EU spending into NHS funding by £350m-a-week. Painted on the side of Vote Leave's big red bus, this pledge was hardly realisable, based on fraudulent data (net EU contribution was merely 160m-a-week at the time), and readily disowned by Brexiters within a few hours after the election was won².

The same year, the US presidential elections saw a similar upsurge of lies in political debates and campaigns. For Donald Trump's in particular, lying has been a fundamental campaigning strategy. Journalist Maria Konnikova went as far as claiming that “the sheer frequency, spontaneity and seeming irrelevance of his lies have no precedent”³. This is hardly an exaggeration: a study shows that (out of a given sample) an estimate of 70% of Trump's statements during the presidential campaign were false, another reported that in a one-hour-long TV appearance Trump managed to utter an astounding total of 71 lies⁴. Trump's attitude has not changed since: this 23rd of June, six months after the elections, his dedication to lying has prompted the *New York Times* to dedicate a full page of the newspaper to print every lie that Trump has publicly told since taking office⁵.

This apparent increase in political lying and deception has brought some commentators to question whether the UK referendum and the US elections were based on informed voting, an essential ingredient to a functioning democracy. Similarly, the legitimacy of the political decisions taken on the basis of such voting has been repeatedly put to question. More generally, these events have brought to the fore the importance of understanding lying and other form of deceptive communication.

¹ Luke Lythgoe and Hugo Dixon, “EU-bashing stories are misleading voters – here are eight of the most toxic tales”, *The Guardian* (Thursday 19 May 2016).

² Ashley Kirk, “EU referendum: The claims that won it for Brexit, fact checked”, *The Telegraph* (13-03-2017); Jon Stone, *The Independent*, “Nigel Farage backtracks on Leave campaign's '£350m for the NHS' pledge hours after result” (24 June 2016).

³ Maria Konnikova, “Trump's Lies vs. Your Brain”, *Politico Magazine* (January 2017); “Donald Trump's File at *Politifact* (url: <http://www.politifact.com/personalities/donald-trump/>);

⁴ Dana Liebelson, Jennifer Bendery, Sam Stein, “Donald Trump Made Up Stuff 71 Times in an Hour”, *Huffington Post* (30-6-2017).

⁵ David Leonhardt & Stuart A. Thompson, “Trump's Lies”, *New York Times* (23-6-2017).

If understanding lies is important, it is not because of politics alone. Lies have been at the centre of the public debate for a number of other reasons. Famous sportsmen have infamously lied about taking performance-enhancing drugs, leading many supporters to question how fair these competitions are: if we consider the case of road cycling alone, famous are the cases of Tour de France winners Marco Pantani and Lance Armstrong, the latter of which was eventually stripped of all his titles. The job of a scientist is the pursuit of truth, but even academics are susceptible to the temptations of lying. One famous case is that of physicist Jan Hendrik Schön. His purported discoveries about molecular semiconductors were so revolutionary that the scientific world was expecting him to win the next Nobel Prize in Physics – then his data turned out to be fraudulent, Schön was fired and his PhD revoked⁶. More generally, humans are prone to lie, way more than we are usually willing to admit to ourselves: studies show that people on average report to lie one or two times a day (DePaulo et al. 1996) – but this rate can go as high as a mean of *three lies a minute* when we consider interactions with strangers (Feldman, Forrest & Happ 2002).

In light of these observations, is not surprising that the study of lying, insincerity and deception is gaining centre stage in studies of linguistic communication: disciplines as diverse as sociology, forensics, psychology, and neuroscience have displayed an increasing interest in their analysis (Levine 2014). Philosophy certainly cannot aspire to develop lie-detection techniques, and much less a ‘cure’ for lying, but it can certainly aim to understand this phenomenon, bring conceptual clarity to its study, and open the way for further investigation. Such is the purpose of the first half of this dissertation.

To give an idea of what kinds of philosophical issues arise with regard to lying, it will be helpful to have a quick overview of the main philosophical debates concerning the issue. Perhaps the oldest recorded philosophical interest in lying can be traced back to the sixth century B.C.: it is in these times that Epimenides of Knossos developed the ‘liar paradox’, challenging logicians to determine the truth value of statements of the form “I always lie” for the centuries to come.

Throughout history, however, the most discussed issue in the philosophy of lying has almost certainly been that of the morality of lying. Questions like “is it ever morally permissible to lie – and if so, under which conditions?” have gripped philosophers throughout history: from Augustine to Aquinas, from Kant to Grotius, many established moral philosophers

⁶ Leonard Cassuto, “Big trouble in the world of ‘Big Physics’”, *The Guardian* (18-9-2002).

have tried to tackle these conundrums. A related philosophical issue is that of the permissibility of political lying. A tradition from Plato to Leo Strauss has defended the claim that lying for the common good is permissible and recommended for a virtuous leader; an opposing tradition disagrees – amongst other reasons, on the ground that political lying undermines democratic institutions (cf. Carson forthcoming). More recently, philosophical interest in lying has emerged in the contemporary debate in epistemology. Around the 1980s, philosophers got more and more interested in the transmission of knowledge via testimony – *i.e.*, knowledge that epistemic agents learn from words. Within this field of inquiry, lies are philosophically significant, as they represent a potential hindrance to the social process of knowledge transmission.

Each of these philosophical questions (except for the liar paradox) presuppose a clarification of the concept of lying: to be able to tackle them, we need first to determine what lying is – and this, in itself, is a philosophical question. Rather than delving into any specific debate on lying, thus, the first part of this dissertation will be devoted to respond to this foundational question: “what is lying?”. More specifically, it will attempt to identify necessary and sufficient conditions for a statement to count as a lie. This apparently easy task will turn out to be one of the most challenging for the philosophy of lying: it is not easy to identify a definition that is resistant to counterexamples; furthermore, an informative analysis of lying demands further analysis of a number of neighbouring concepts – such as insincerity, assertion and deception.

In attempting to define a concept, this dissertation places itself in the tradition of conceptual analysis. A central idea within this tradition is that we can learn something about a given concept by breaking it down into its essential components. In this sense, the desideratum of a theory of lying is to identify the necessary and sufficient conditions for something to be a lie.

From a methodological point of view, this means that a good analysis of lying should neither be too broad (it should not be subject to counterexample to the sufficiency of the analysis) nor too narrow (there should not be counterexamples to its necessity). The litmus test for a good definition here are intuitions: there is consensus that a good definition should reflect our intuitions about particular cases. There is disagreement as to the scope of ‘our’ here – some scholars think that a good definition should reflect laypeople’s intuitions, other that expert intuitions are more important. Ideally, a good definition should meet the intuitions of both groups.

As I proceed in my analysis, I will argue in favour of a number of claims about lies. Some will be claims about conceptual possibilities and unusual lies: that there can be lies that are not intended to deceive; that there can be lies that are true, that there can be lies that we believe to be true. Even though each of these claims may sound astounding, I will show that they are perfectly consistent with our intuitions about particular cases, and relevant to the understanding of what lying is. Some other claims will be about the inability of traditional definitions to explain some standard features of lying: for instance, standard definitions struggle to explain that one can lie by promising something, that all lies always involve the undertaking of some distinctive responsibility, and that lies can be more or less insincere, depending on our degree of confidence of our beliefs, and on the hedges and other linguistic devices that we use to express such beliefs. As I review and criticise the existing literature, I will develop my own definition of lying – hopefully, one that will be useful to solve other philosophical problems about lying, such as ethical and epistemological ones. For a more detailed breakdown of the topics that will be touched in this part of thesis, the reader can refer to the *Plan of the Work* (I.3).

2. Incorrect Assertions

While the first part of the thesis deals with insincere communication, the second part deals with irresponsible communication: statements that are at fault not because they are insincere, but rather because they are false, or because they are not supported by adequate evidence. There are uncountable real-world examples of this sort of communicative impropriety; to find one to which we are all acquainted, it will be helpful to consider some popular conspiracy theories. Even if we limit our attention to wildly implausible (yet popular) ones, the list of such theories is endless: one claims that the Earth is hollow and secret civilisations live inside it, in the hidden city of Agartha, illuminated by an internal sun; another that the Earth is flat, and that governments *create* false evidence to trick us into believing the opposite; others that our planet has been visited by aliens, and governments operate to *destroy* all signs of contact with extra-terrestrial life.

What is at issue here is not why conspiracy theorists believe in these theories, or what is wrong with such beliefs. It is rather the role that conspiracy theorists have in the propagation

of these falsehoods – their role as communicative agents. Conspiracy theorists are not liars⁷: when they assert that the Earth is flat or hollow, they do it in good faith, sincerely. But even if their assertions are not lies, they are importantly at fault.

In the age of social media, this form of irresponsible communication has become even more problematic and apparent. On the one hand, social media make it easier for people to share false information, without verifying its veracity. On the other hand, social media allow this information, shared in good faith, to reach a wider audience – often with catastrophic effects. One infamous example is the epidemic sharing of the fake news that Pope Francis endorsed Donald Trump during the US electoral campaign. Initially posted by the satirical fake news website WTOE 5 News, the news was shared by roughly a million of Facebook users, giving the rumour a resonating voice: it was the most shared news posted during the elections, outperforming any real one at the moment. Considering a larger sample, a study shows that pro-Trump fake news stories received around 30 million shares on Facebook during the electoral campaign, and often more attention than real ones⁸. Overall, this kind of data strongly suggests that the irresponsible communicative behaviour of some voters (their sharing unconfirmed news coming from dubious sources) might have strongly influenced the outcome of the last US elections.

The communication of false information and unwarranted claims thus poses a serious threat to a citizen’s access to reliable, factual information – and to some extent to democracy itself, as long as we understand this institution to be reliant on informed voting. In some cases, it can even be a threat to people’s lives and wellbeing. Such in the case of the disinformation spread by the anti-vaccine movement, that has led to an upsurge in deaths by otherwise curable diseases, with an estimated death toll of about 9 thousand people since 2007⁹.

I have already pointed out that conspiracy theorists and anti-vaccine activists are not lying when they propagate their falsehoods (since they believe in them), and that nonetheless their communicative behaviour is at fault. If not lying, which kind of communicative fault is committed in these cases?⁹ The second part of the dissertation attempts to answer this

⁷ An exception might apply to the initial propagator(s) of the theory, who could have invented it with deceptive intentions. For the purpose of the present discussion, let us call a *conspiracy theorist* only someone who *genuinely believes* in a false and implausible conspiracy theory.

⁸ Craig Silverman, “This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook”, *Buzzfeed* (16-11-2016); Jason Tanz, “Journalism fights for survival in the post-truth era”, *Wired*, (14-2-2017).

⁹ Anti-vaccine body count, at <http://www.jennymccarthybodycount.com/>

question: it aims to understand, on a general level, what kind of communicative expectation is violated when people assert false or unwarranted propositions. This sort of question has gained centre stage in the philosophical debate in recent times. For reasons that I will discuss more in detail later, in the last 20 years philosophers have become more and more convinced that assertions (claims that something is true) are governed by a single rule – a rule whose violation can explain the distinctive wrongness involved in uninformed and false statements. Scholars working in this tradition tend to adopt the following hypothesis (or a hypothesis along the following lines):

Assertion is the only speech act governed by the following rule: you should not assert a given proposition unless condition C is satisfied

According to this hypothesis, there is a condition that must be satisfied for an assertion to be appropriate, and if we replace ‘C’ with that condition, we obtain the norm that governs all assertions. This hypothesis offers a simple model to explain what is wrong with assertions by uninformed speakers (like conspiracy theorists and anti-vaccine activists): these assertions are at fault not because they are lies, but because they violate the norm governing all assertions. Just like we expect people not to lie, we expect speakers to follow the norm; when they fail to do so, a violation occurs, making ground for reproach, criticism and blame.

Given that the hypothesis leaves condition C unspecified, a great deal of the explanation will turn on how one specifies condition C. To see a few candidates, let us consider a real life scenario. Suppose that, after reading about the fake news that the Pope endorsed Trump on the satirical news site WTOE 5 News, an American citizen (call him Ferdinand) tells another friend that the pope endorsed Trump as a presidential candidate. Now, we might have the intuition that Ferdinand’s assertion is wrong because it is not true, which gives us a first candidate for condition C: do not assert a proposition *unless it is true*. But some other people might have the intuition that Ferdinand’s assertion is wrong for different reasons. For instance, Ferdinand does not have the appropriate evidence to make such a claim: he failed to check whether the information came from a reliable source (as a matter of fact, it did not), which is especially pernicious given the implausibility of the proposition (there are many known ideological disagreements between the Pope and Donald Trump). If one has this intuition, it is not the falsity of the proposition that should explain the wrongness of Ferdinand’s assertion; rather, the lack of appropriate evidence. We now have a second candidate formulation of condition C: do not assert a proposition *unless you have*

appropriate evidence. Yet, some other people might intuit that Ferdinand's assertion is at fault in both ways: you should neither assert what is false, nor what is not supported by appropriate evidence. If this is the case, then perhaps condition C is that a speaker should *only assert what he knows* (as knowledge requires both truth and appropriate evidence).

These are but a few possible solutions to the question raised by the hypothesis. There is a vast literature on the issue, ripe with different accounts of what makes an assertion permissible. For the moment, I will not enter the intricacies of this debate, nor will I attempt to explain which account offers a better explanation of the norm regulating assertion. All I aimed to clarify is that the hypothesis that assertion is subject to a single epistemic norm has the potential to explain what is wrong with assertions that are not lies, but that are nonetheless false or unwarranted. This will be the explanandum of the second part of the dissertation. In those pages, I will assess the solidity of the hypothesis to which the scholars involved in the debate subscribe, and attempt to establish which kind of condition C should be.

3. Outline of the Dissertation

As already mentioned, this dissertation is divided into two parts. **Part A** discusses problems concerning lying, and **part B** problems relating to the norm of assertion.

Part A opens by introducing the philosophical debate about the definition of lying. In the contemporary discussion, two factions are opposed: *deceptionist*, who believe that lies necessarily aim at deceiving someone, and *non-deceptionist*, who deny this necessary condition. **Chapter 2** presents some known arguments against deceptionist definitions, and supplements them with novel arguments. It concludes by responding to some recent objections presented by deceptionists in reply to these arguments. **Chapter 3** turns to criticisms of the other faction, non-deceptionism. It opens by recapitulating some known counterexamples that affect each of the most popular non-deceptionist definitions. It then proceeds to show that all these accounts have a common defect in common: they cannot deal with three 'speech-act theoretic' puzzles about lying - puzzles that involve lies performed by means of a speech act other than assertion. To find a solution to these puzzles, I turn to speech act theory, with the aid of which I develop a non-deceptionist definition that avoids the objections to which all the other definitions are subject. In order to further refine this account, **Chapter 4** tackles the problem of defining insincerity. It is divided into four main sections. The first extends the insincerity conditions to speech acts

other than assertion, showing how insincerity is not necessarily a matter of beliefs. The second tests this account against ordinary speaker intuitions, finding support for the proposed account. The third deals with the problem of graded insincerity. Speakers can be more or less confident in what they say, and communicate a higher or lower degree of confidence in what they say. Insincerity involves some discrepancy between the belief held by the speaker and the belief communicated by its utterance: my account proposes a system to model this issue and define which degree of discrepancy is required to call something a lie. In the closing section, I bring together all the findings, to develop a unified definition of lying.

Part B begins by introducing the norm of assertion hypothesis. According to this hypothesis, there is one norm that defines whether a proposition can be asserted or not, and this norm takes the form: “assert that p only if p has C ”. In order to identify which propositions are assertable, philosophers have to determine exactly which property is C . **Chapter 5** puts some pressure on this hypothesis, and more specifically on the idea that the envisaged norm is *constitutive* of assertion. It shows that different authors writing on the norm of assertion have interpreted this claim in different ways. I argue that no matter which interpretation is chosen, the idea that the envisaged norm is constitutive of assertion is misguided. Once one recognises this, the assumption that assertion is subject to *only one* norm of this kind loses its appeal, and so does the idea that assertion can be defined as the only speech act that is only subject to this rule. **Chapter 6** reviews two ways of fleshing out the norm of assertion hypothesis: *factive* and *non-factive* accounts. The former maintain that only true propositions are assertable, whereas the latter deny this, and argue that some false propositions are assertable. Factivist positions are subject to known counterexamples (*unlucky assertions*), but they are generally defended with a reasonable argument: that only a factive account can explain why false assertions are incorrect and liable to criticism. I deny this claim, showing that the same data can be explained by taking truth to be the *purported aim*, rather than the rule, of assertion. I conclude by arguing that, in the light of this data, the account offering the best prediction will be one featuring one or more non-factivist rules, paired with the view that assertion aims at truth. In **Chapter 7** I summarise the key ideas that I defended throughout the thesis.

Part A- Lying

A. The definition of lying

Each of the two parts of this dissertation will feature an introductory chapter like the present one – labelled with a letter, instead of a number, to indicate its different nature. In this introductory chapter, the philosophical debate on the definition of lying is presented. After a short historical introduction, I introduce the *classic definition* of lying. This definition involves two key conditions: the *statement condition* and the *insincerity condition*. Each is briefly introduced and explained with the help of examples. I then proceed to consider the contemporary debate, that involves an opposition between two opposite factions: *deceptionist* and *non-deceptionist* (or *assertion-based*) definitions. After presenting each of them, I set the ground for the discussion in the further chapters.

1. The Classic Definition of Lying

As Augustine gracefully points out in his *De Mendacio* (possibly the first work to systematically discuss this issue), the question of lying “is, indeed, very full of dark corners, and has many cavern-like windings, whereby it oft eludes the eagerness of the seeker; so that at one moment what was found seems to slip out of one's hands, and anon comes to light again, and then is once more lost to sight”. Augustine is stressing an important truth: defining lying (and the morality thereof) is a more difficult task than it seems at first sight. As a matter of fact, the history of philosophy has proved him right: in almost two millennia of discussion, scholars have not reached a consensus about which definition best captures the concept of lying.

Nevertheless, even though there is no agreed upon definition of lying, there are some firm points of agreement. If our discussion of the definition of lying has to start somewhere, it probably has to start from such firm points. Three in particular seem to be resistant to even the strongest scepticism. First, lying is an intentional act: there is no such thing as lying unintentionally. Second, lying requires making a statement: in order to tell a lie, you have to produce some linguistic token. Third, lying requires making an insincere statement: there is no such a thing as a sincere lie, because lying involves stating something that you do not believe.

From these three firm points, we can get to a first tentative definition of lying:

Classic definition:

(CD) To lie is to intentionally state something that you believe to be false

This definition gets most simple cases right. Consider an obvious example of lying: if Pinocchio tells Geppetto “I went to school this morning” even if he did not go, intuitively he has lied. Since Pinocchio is intentionally telling Geppetto what he believes to be false, the definition correctly dubs this case a lie. Now, suppose that later that day, Pinocchio is asked where Geppetto is. Pinocchio has seen Geppetto at the workshop, so he replies “My dad is at the workshop”. It turns out that Pinocchio is wrong, as Geppetto snuck out of the workshop to visit his lover in secret. In this case, Pinocchio has said something false, but his statement is sincere: intuitively, he has not lied. Once again, the intuition is tracked by the definition, that rules out this case as the statement is not believed to be false. At first blush, the classic definition deals correctly with simple cases of lying, and correctly distinguishes them from other false or deceptive utterances.

I called this definition ‘classic’. This term may seem unorthodox to some readers, as philosophers (e.g. Lackey 2013, Mahon 2015) often have in mind a different definition when they talk about the *standard* or *traditional* definition of lying (which I will soon introduce as the ‘deceptionist definition’). I purposefully chose the term ‘classic’ instead of ‘standard’ or ‘traditional’ to acknowledge this departure from orthodox terminology. Such departure is justified by two reasons. The first is that this was very likely the first philosophical definition to appear in the literature, and remained the standard definition for at least a thousand years: introduced by Augustine (DM, IV AD), it is still considered the standard view in the works of Aquinas (ST, XIII AD) and Peter Lombardus (SEN, XIII AD)¹⁰. The second reason is that, while there is no consensus as how to define lying, the overwhelming majority of scholars agree that intentionally stating what you believe to be false is a *necessary* condition for lying: you cannot lie unless you meet the conditions

¹⁰ Contrary to what is argued by some authors (Siegler 1969: 129n), Aquinas did not require a deceptive intention: “The desire to deceive belongs to the perfection of lying, but not to its species, as neither does any effect belong to the species of its cause” (Aquinas ST, II, IIae, q110, a2). Augustine’s case is more complex: the traditional interpretation is that he endorsed the intention to deceive condition (e.g. Siegler 1969: 129n; Feehan 1988: 135-8) but according to Griffiths (2004: 30) this view is misguided, and there are better reasons to think that he was partisan of a “non-deceptionist” definition of lying. I believe that an accurate and charitable reading of Augustine’s work can only support an interpretation lying between these two extremes: Augustine simply did not settle the question as to whether a deceptive intention is required for lying or not.

stated by (CD). In other words, the current consensus is that (CD) offers an accurate *characterisation* of lying, even if it is arguably not an accurate enough *definition* – because meeting the conditions stated in (CD) may not be sufficient for a statement to qualify as a lie.

The current debate on the definition of lying mainly revolves around the question of which additional condition is best suited to make the definition meet sufficiency. Before entering the current debate, however, it is worth familiarising ourselves a bit more with the points of agreement set by the classic definition. More specifically, in the next two sections I will discuss in detail the two main requirements posited by the classic definition: the statement condition and the (intentional) insincerity condition.

1.1 The Statement Condition

You! You chameleon! Bottomless bag of tricks! Here in your own country would you not give your stratagems a rest or stop your spellbinding for an instant?

Homer, *Odyssey*, XXIV

In his *Parerga e Paralipomena* (1851/1974:538), Arthur Schopenhauer writes that “there is in the world only one mendacious and hypocritical being, namely man. Every other is true and sincere, in that it frankly and openly declares itself to be what it is and expresses itself as it feels”. Unlike animals, whose ingenuity and transparency is fascinating to us, the degenerate human tendency to lie “stands as a blot on Nature”. Schopenhauer’s severe verdict is certainly inaccurate, given that plants and animals are capable of incredibly complex forms of deception. However, in his observation we can find a grain of truth: arguably, non-humans cannot lie, at least if lying requires *telling* a lie – that is, uttering a linguistic token that we believe to be false.

The intuition that lying requires linguistic abilities is reflected by the classic definition of lying, according to which lying involves stating a proposition, or saying something. Following Mahon (2015), I call this the *statement condition* for lying. The statement condition captures the intuitive difference between lying and other forms of deception. Pace to Schopenhauer, countless examples of deception can be found in nature to illustrate this distinction. For instance, the orchid *Cryptostylis erecta* is pollinated by the so-called orchid dupe wasp (*Lissopimpla excelsa*), the males of which mistake the flower parts for female

wasps, and copulate with them. While there is a sense in which the wasp is *deceived* by the orchid's shape, it would be erroneous to say that the orchid *lied* to the wasp. The statement condition thus captures the intuitive distinction between lying and simple deception.

A few authors reject the difference between deception and lying¹¹. However, this leads them to conclusions that are rather counterintuitive. As a *reduction ad absurdum*, it will suffice to remark that Smith (2004), who rejects the statement condition, lists as lies “breast implants, hairpieces, fake orgasms and phony smiles, as well as age-concealing make up and deodorants that disguise our scent”. While these cases usually involve attempted deception, they are clearly not lies. Claiming that wearing deodorant is lying can strike as bizarre, and illustrates the counterintuitive consequences of rejecting the statement condition.

The view that lying involves stating something can be traced back to Augustine's seminal work. In his *Contra Mendacium* (XII) he writes that “a lie is a false signification by words”. Aquinas (ST, q110), who knew Augustine's work thoroughly, rightly specifies that this does not mean that lying is necessarily a matter of verbal communication:

As Augustine says [DDC, III], words hold the chief place among other signs. And so when it is said that "a lie is a false signification by words," the term "words" denotes every kind of sign. Wherefore if a person intended to signify something false by means of signs, he would not be excused from lying

These observations are the right track: there are cases of lying that do not involve the use of words. One can lie by using any sort of conventional signals, or combinations thereof. For instance, you can lie by using body gestures that have conventional meaning (as in nodding with your head to agree), or using smoke signals, and so on. To capture all these cases, scholars working on lying usually refer to Chisholm and Feehan's (1977:150) definition of statement:

¹¹ These are generally biologists (e.g. Smith 2000, Dawkins 1989:64) and psychologists (e.g. Ekman 1985:26-8, Vrij 2008), who often treat the verb “lying” as equivalent to “intentionally deceiving”.

Definition of Statement (Chisholm & Feehan)

S states that p to A iff

(a) S believes that there is an expression E and a language L such that one of the standard uses of E in L is that of expressing the proposition p;

(b) S utters E with the intention of causing A to believe that he, S, intended to utter E in that standard use

In other words, a lie has to be a linguistic token that is believed to express a proposition, and that is believed to do so in virtue of some linguistic convention. This formulation of the statement condition is preferable to Augustine's, as it does not seem that "every sign" used to signify something false can be used for lying. Consider the example of wearing a lab coat to pretend that you are a scientist, or a ring to pretend that you are married. In these cases, you use a sign to communicate something false, but you are not lying. The statement condition proposed by Chisholm and Feehan correctly rules out these cases. To go back to Schopenhauer's erroneous claims about nature's intrinsic sincerity, another example of deceptive signals that are not lies is found in animal signalling. Many animals produce signals that are associated with a stimulus. Sometimes, however, they use such signals in the absence of the stimulus, for deceptive purposes. For instance, *Lanio Versicolor* sentinel birds are known to produce alarm calls in the absence of predatory birds, in order to scare their conspecifics when the competition for food is high (Munn 1986). According to the statement condition, and consistently with intuitions, these alarm calls are deceptive, but are not lies¹².

Lastly, some authors (Siegler 1966, MacCormick 1983, Fallis 2010; 2013; 2014, Meibauer 2011; 2014, Saul 2012, Stokke 2013; 2017, Viebahn 2017) prefer to use the term 'saying'

¹² Some might argue that some animal signalling can count as lying. It is unclear whether non-human animals can meet the statement condition, which requires the ability to communicate in a language that assigns meanings to expressions. But if some animals can communicate in a language, it could be argued that they can also lie. Some cases of animals that purportedly used language to lie are reported in the literature. For instance, Fouts & Mills (1997:156) report the following dialogue with Lucy, a chimpanzee trained to speak in American Sign Language:

Fouts: What that? [indicating a pile of chimpanzee feces on the floor]

Lucy: What that?

Fouts: You know. What that?

Lucy: Dirty dirty.

Fouts: Whose dirty dirty?

Lucy: Sue. [a reference to Sue Savage-Rumbaugh, a graduate student of Fouts]

Fouts: It not Sue. Whose that?

Lucy: Roger!

Fouts: No! Not mine. Whose?

Lucy: Lucy dirty dirty. Sorry Lucy.

in place of ‘stating’. In the literature, the two terms are understood to be synonyms: they both indicate the utterance of a meaningful declarative sentence. However, it should be noted that a minority of authors that use the term ‘saying’ (Saul 2012, Viebahn 2017) do this to mark their commitment to a slightly different version of the statement condition – more specifically, a Gricean account of what it means to *say* something. For the purpose of this dissertation, however, we can safely ignore this subtle distinction. Consequently, in this dissertation I will treat ‘saying’ and ‘stating’ as synonymous.

1.2 The Insincerity Condition

*Non enim omnis qui falsum dicit mentitur
si credit aut opinatur verum esse quod dicit*

Augustine, *De Mendacio*, 3.3

According to the insincerity condition, the speaker has to *believe* that what is said is false. I call this condition the *insincerity* condition, because it captures the difference between an insincere utterance (that you believe to be false) and a mere mistake (that you believe to be true, but turns out to be false). Going back to the previous example, if Pinocchio mistakenly believes that Geppetto is in the workshop and voices his mistake, he says something false, but he does not lie: while incorrect, his utterance is sincere. The insincerity condition captures this intuition: saying something false does not amount to lying, unless the falsehood is also believed to be false.

Importantly, requiring that the statement is believed to be false is not yet requiring that the lie must be false. Most scholars deny that actual falsity is necessary for lying¹³. However, a few authors (Grotius RWP:1209, Benton forth, Turri & Turri 2016) have suggested that falsity of the statement, in addition to *belief* in its falsity, is required for lying. For these authors, if your believed-false utterance turns out to be true, you have not lied. People seem

¹³ Augustine (DM, 3.3), Aquinas (ST, II-II, q.110, a1), Kant (1797), Leonard (1959:182), Isenberg (1964:466; 1974), Lindley (1971), Mannison (1969:138), Chisholm & Feehan (1977), Kupfer (1982:104), Adler (1997), Williams (2002), Mahon (2008), Fallis (2009). Other authors endorse weaker positions. Carson (1982:16; 2006:284; 2010:39) and Saul (2012) provide a definition that does not require falsity, but both suggest that the definition could be strengthened to include such a requirement. Siegler (1966:132) suggests that falsity is necessary for telling a lie but not for lying; less controversially, Coleman & Kay (1981:28) argue that falsity is necessary for *prototypical* lying.

to have different intuitions in this respect. For our purposes, it will suffice to say that the classic definition gives the opposite verdict, and counts inadvertent truth-telling as lying. The insincerity condition also grounds the intuitive distinction between lying and merely misleading. A classic example of misleading that falls short of lying is found in the hagiography of St. Athanasius (MacIntyre 1994: 336):

Persecutors, dispatched by the emperor Julian, were pursuing [Athanasius] up the Nile. They came on him travelling downstream, failed to recognise him, and enquired of him: "Is Athanasius close at hand?" He replied: "He is not far from here." The persecutors hurried on and Athanasius thus successfully evaded them without telling a lie.

In this example, Athanasius is attempting to deceive his pursuers by implying that he is not Athanasius. While deceptive, his carefully phrased sentence is true; intuitively, it does not constitute a lie. The insincerity condition makes sense of this intuition: it predicts that this is not a lie, given that Athanasius does not believe that what he has said is false.

Finally, according to the classic definition the speaker has to *intentionally* say what he believes to be false. This requirement of intentionality is found in some classic texts; here is how Griffiths (2004:27) introduces it:

There is an internal fact and an external fact. The internal is what's in the mind (animus—Augustine often also uses "heart," cor, or "mind," mens, for the same purposes) and the external is what's said or communicated in some other way—by gesture or expression or some other nonverbal sign. Lying happens when the two are intentionally separated. (my emphasis)

One way of understanding the requirement that the speaker *intentionally* says what he believes to be false is for it to be a refinement, or specification, of the insincerity condition. To be insincere, in this sense, is to *intentionally* establish a discrepancy between what you believe and what you say. To some, the adverb 'intentionally' may seem redundant here. After all, unintentional falsities (*i.e.* mistaken but sincere statements, as in Pinocchio's statement that Geppetto is in the workshop) are already ruled out by the requirement that the speaker says something he *believes* to be false. However, the intentionality constraint plays an important role here: it rules out other species of accidental falsehoods from the definition of lying and insincerity.

Linguistic mistakes (cases in which when the speaker misspeaks, or gets confused about the meaning of words) offer a first example of such accidental falsehoods. Saul (2012:14, see also 2012:16) has a rather amusing example:

Anna, an English rock climber, wanted to tell her [Mexican] colleagues that many people in England climb without ropes. So she uttered (2):

(2) En Inglaterra hay mucha gente que escala sin ropa

(2) actually means that in England there are many people who climb without clothes [ropa]. This claim is false, but Anna did not lie; she accidentally said something false, through a linguistic error.

A definition that only requires a liar to state a believed-false statement would count (2) as a lie: Anna has said that people in England climb without clothes, and she believes this to be false. However, Anna has not said what she believed to be false *intentionally*, so that this case is not counted as a lie by the classic definition. The intentionality requirement is thus indispensable to maintain the distinction between lies and malapropisms.

Self-deception also generates examples of accidental falsehoods that can only be addressed by the intentionality constraint on insincerity. Self-deception is a psychological condition whereby one convinces oneself of the truth of something that one knows to be false (hence the deception), and has no awareness of being so deluded. In other words, when you are self-deceived about p , you believe that you believe that p , but you do not in fact believe that p . Ridge (2006:488–9) offers an example:

Bob believes that he believes his mother loves him but actually does not believe that she loves him. In fact, Bob believes his mother hates him. [...] Suppose we ask Bob whether his mother loves him and he says, “Yes, of course she does”.

In this case, Bob is saying something that he believes to be false. Nonetheless, he is not saying something he believes to be false *intentionally*. Once more, the intentionality constraint on insincerity is doing the heavy lifting in distinguishing lying from other accidental falsehoods: were we not to require that the “separation between mind and words” was intentional, we would incorrectly count these cases as lies.

2. The Current Debate: Deceptionist vs Non-Deceptionist accounts

There should be no doubt at this point as to the exact meaning of (CD) within the debate on lying, and the arguments that support it should be clear.

Classic definition:

(CD) To lie is to intentionally state something that you believe to be false

Now, I have already mentioned that while there is consensus that (CD) offers an accurate *characterisation* of lying (it is correct about what is required for lying), it is not taken to be an accurate *definition*, because meeting the conditions stated in (CD) is not sufficient for a statement to qualify as a lie.

This criticism of the classic definition is grounded: there are indeed statements that are believed to be false and are not lies. Believed-false statements that are not lies include *ironic statements; fictional statements* (e.g. uttered on stage or written in a fictional novel); *jokes; teasing remarks; hyperboles; metaphors; euphemisms*, and so forth. These kinds of statements, that I will call *non-assertoric falsehoods*, represent solid counterexamples to the classic definition: they meet (CD), but intuitively are not lies.

To see this, let us consider an example of non-assertoric falsehood, a fictional statement. Imagine that an actor on stage utters:

(1) I am Ubu, Prince of Podolie, Duke of Courlande, Earldom of Sandomir, Margrave of Thorn

The actor is saying something that he believes to be false, but he is clearly not lying – he is just pretending to be Ubu for the sake of the play. Since the classic definition has no resources to rule out these cases, it must be incorrect. To be sure, this is quite an important failure: it means that the definition gives incorrect verdicts in a very wide variety of cases, involving virtually every figure of speech that can be literally false.

Upon the failure of the classic definition, there are essentially two strains of definitions that are able to solve this problem and distinguish between lies and other *non-assertoric falsehoods* as non-lies: *deceptionist definitions* and *non-deceptionist, assertion-based definitions*. Deceptionist definitions expand (CD) by introducing the further condition that the speaker must intend to deceive his audience. This amendment deals with *non-assertoric falsehoods* by excluding them in virtue of the fact that they are not intended to deceive. For

instance, the actor's utterance that (1) is not counted as a lie, because the actor is not attempting to deceive his audience. Broadly, deceptionist definitions¹⁴ are phrased as follows:

Deceptionist definitions:

S lies to A iff:

(a) *S states that p*

(b) *S believes $\neg p$*

(c) *S intends A to believe p*

In recent years, there has been growing consensus that these definitions are incorrect. Their key problem is that the 'intention to deceive condition' (c) exposes the definition to several counterexamples (that will be discussed extensively in the next chapter).

Are there alternative ways to amend the classic definition? The most influential alternative is to require that the speaker *genuinely asserts* that *p*. This amendment deals with *non-assertoric falsehoods* by excluding them in virtue of the fact that they are not genuinely asserted. For instance, the actor's utterance on stage does not count as a lie because the actor is not genuinely claiming that he is King Ubu, but merely pretending to claim for the sake of the play. Definitions that follow this strategy are dubbed *non-deceptionist*, because they reject the intention to deceive conditions, and *assertion-based* - because, unlike the classical definition (which is also *non-deceptionist*), they introduce the further requirement that the relevant proposition is asserted. More formally, assertion-based definitions read:

Assertion-based (non-deceptionist) definitions:

S lies to A iff:

(a) *S says p*

(b) *S believes $\neg p$*

(c') *S asserts that p*

In sum, the contemporary debate on the definition of lying articulates around which putative additional condition is required to amend the classic definition of lying: some authors believe that the speaker has to *intend to deceive* the audience, other that he has to genuinely *assert* the relevant proposition. Chapter II will deal with the deceptionist

¹⁴ This label can itself be deceptive, as it may be interpreted as suggesting that these accounts require successful deception. Only intended deception is required: "deceptionist" should be taken to be a shorthand for "based on the intent to deceive condition".

accounts, and Chapter III will discuss their most prominent alternative, assertion-based accounts.

IV. Insincerity

Sincerity is often valued as an important virtue, and insincerity criticised as a vice. We generally trust other people to be sincere, and their testimony is a fundamental source of information without which we could hardly get on with our ordinary lives. For these reasons, insincerity has elicited the interest of philosophers working not only on language, but also on ethics and epistemology. Any epistemological or ethical discussion of insincerity, however, presupposes settlement of one fundamental question: what is it to be insincere? This chapter is devoted to answering this question. More specifically, it is concerned with two related aims. Its primary aim is to provide an analysis of insincerity as a component of lying. In other words, the primary goal of this chapter is to refine the insincerity condition in the definition of lying, specifically in the light of some counterexamples to which the traditional definition falls victim, including the third speech-act theoretic puzzle introduced in the previous chapter. But even if it is the interest in lying that drives this enquiry, my analysis of insincerity as a necessary condition for lying arguably retains value also as an analysis of insincerity in general. My related, secondary aim is thus to provide a characterisation of insincerity that (at least to some extent) applies also outside the debate on defining lying.

This chapter is divided in four long⁵³ sections. The first one simply lays out the problem of defining insincerity, and clarifies more accurately which notion of insincerity I am after. In section 2, I attack the *only-belief view*, namely the idea that beliefs are the only attitude that is relevant to determining whether an utterance is a lie. As I have argued in the previous chapter, this account of the insincerity condition for lying is inaccurate: for instance, a promise can be a lie when the speaker *does not intend* to perform the promised action. I consequently expand the insincerity condition to attitudes other than belief, thereby solving the last of my three speech-act theoretic puzzles. In the section 3 I test my revised condition empirically, showing that ordinary speakers share the intuition that also intentions can determine whether an utterance is a lie.

⁵³ The longer sections in this chapter are based on material that I have already published elsewhere. Section 2 draws on Marsili (2016); section 3 is an almost literal excerpt from the same paper. Section 4 is based on material from two different papers, Marsili (2014) and Marsili (2017).

Section 4 criticises the ‘dichotomic view’ of insincerity, namely the idea that either a statement is believed to be true, or believed to be false. It introduces lies that fall outside this dichotomy, namely *fuzzy-lies* and *graded-belief lies*, and develops a refined insincerity condition that treat these graded lies correctly: broadly, a speaker is insincere if he believes his statement to be more likely to be false than true. The final, fifth section brings together all my findings into a general account of the insincerity conditions for lying.

1. INSINCERITY: A PRELIMINARY ACCOUNT

In ordinary language, the terms ‘sincere’ and ‘insincere’ are used in different contexts with different meanings. Before initiating a more thorough discussion of insincerity, I would like to clear up some ambiguities about these different meanings, to explain exactly which notion of insincerity this chapter aims at analysing.

First, in ordinary language insincerity need not refer to linguistic utterances. We can say that a *smile*, or even *person* (as opposed to an utterance) is insincere. While these uses of the term are certainly appropriate, they are not the object of our interest here. This thesis is concerned with insincere utterances and more specifically assertions, and consequently only with linguistic insincerity.

Second, even when we limit our analysis to linguistic insincerity, it seems that this term can be used at least in two ways⁵⁴. In a broad sense, calling an utterance insincere is describing it as *deceptive*, or *aimed at deceiving*. For instance, Bernard Williams (2002: 74) defines insincere assertions as those that “have the aim of misinforming the hearer”. Under this conception, an insincere assertion is one intended to deceive: not only lies, but also *misleading* but literally true statements, *omissions*, and any sort of deceptive statement.

This is not the conception of insincerity I am concerned with, for at least two reasons. First, I have already stated that I am after a notion of insincerity that, paired with a notion of assertion, will provide us with all the notions required for defining of lying. Understanding ‘insincere’ as synonymous to ‘deceptive’ or ‘intended to deceive’ would not help in this enterprise: it would not allow for a distinction between lying and merely misleading (cf. A.1.2), and it would conflate the insincerity condition for lying with the intention to deceive condition. Second, on this conception the analysis of “S was insincere (in saying *p*)” would be equivalent with the analysis “S was deceptive (in saying *p*)”. Between a notion that overlaps with another and one that does not, the latter is clearly more appealing, as it enriches our conceptual toolbox in a way that the former does not.

The sense of insincerity in which I am concerned is then a different one. Under this conception, insincerity indicates a discrepancy between the psychological state of the speaker (*e.g.* believing, intending, desiring) and the psychological state expressed by his speech act (*e.g.* asserting, promising, requiring). Defining ‘insincerity’ amounts to defining the nature of this discrepancy, which will be the subject of this chapter.

⁵⁴ For a review of different conceptions of insincerity, cf. Eriksson (2011).

Finally, insincerity is a complex phenomenon, and there are several philosophical problems that concern this notion, more than this chapter could potentially discuss and analyse. In particular, here I will focus on two problems that have been rarely, if ever, discussed in the literature: how a definition of lying can deal with attitudes other than belief, and with graded insincerity. To focus on these relatively new problems, I will leave aside some classic ones. One in particular will not be discussed in detail: that of differentiating between insincerity on the one hand, and misspeaking and self-deception on the other. I have briefly addressed this problem in A.1.2, where I argued this distinction can be made explicit just by specifying that the speaker has to satisfy (any version of) the insincerity condition *deliberately*. In order to focus on other issues, I will leave further philosophical problems concerning misspeaking and self-deception aside⁵⁵.

⁵⁵ In a recent paper, Jessica Pepp (forthcoming) mentions some further difficulties that may arise in this respect (cf. also Chan & Kahane 2011, Stokke 2014). Considering that Pepp's problems seem to emerge from problems affecting theories of reference in general rather than insincerity in particular, and given that this chapter is concerned with problems that are distinctive of defining lying, I will leave the discussion of these subtle counterexamples for another time.

2. BEYOND BELIEF: INSINCERITY AND ILLOCUTIONARY ACTS

In the previous chapter, I have mentioned that there seems to be universal agreement in the literature on lying that whether an utterance is a lie, and therefore insincere, is only a matter of what the speaker believes. I called this view, that is endorsed by virtually every author in the literature, the ‘only-belief’ view of the insincerity condition for lying.

ONLY-BELIEF: the only attitude (or lack thereof) relevant to determine whether an utterance is a lie is belief

In III.2.3, I have argued that this view is wrong. I will now readdress such criticism more in detail, and show how the insincerity condition for lying can be expanded to other propositional attitudes (e.g. intentions, desires) in order to address this objection. I will develop an alternative account of what it is for a speech act to be sincere or insincere, and then put it to work against counterexamples based on insincere promises. In the next section, I will show that this account better reflects ordinary speaker’s intuitions about what counts as a lie.

2.1 The third speech act theoretic puzzle

Let us start by briefly recapitulating the speech act theoretic puzzle introduced in III.2.3. The example is meant to show that there can be cases of lying in which the speaker believes that the propositional content of the utterance (identified in a non-descriptivist fashion) is true, and in which the speaker’s intentions, rather than his beliefs, are relevant to determine whether his utterance is a lie.

UNFAITHFUL WIFE

Baba and Coco are a married couple. Baba is away from the city for work, and is planning to go out this night. Since Coco is extremely jealous, he asks her: “Will you be cheating on me tonight?”. Baba replies:

(1) Do not worry Coco: I promise that I will not cheat on you tonight

In fact, Baba intends to do her best cheat on Coco at the party, but she is virtually certain that she will end up not doing so, as her terribly awkward manners have always prevented her from seducing any man other than Baba.

In the example, Baba has an insincere intention, and her promise is deceptive: it seems intuitive that (1) is a lie. Nonetheless, the *only-belief* view incorrectly predicts that this is not a lie, because Baba is almost certain that she will not cheat on Coco. What is missing here is an intention (the intention to try to stick to the promise) rather than a belief. Against the predictions of traditional definitions, some utterances can be lies despite their content being believed to be true.

To avoid this counterexample, a definition of lying should allow for propositional attitudes other than belief to determine whether one's utterance is a lie. But how are we to extend the insincerity condition to other attitudes? Speech act theory offers a promising theoretical framework for this purpose: it is a standard view in speech act theory that insincerity can depend on a variety of attitudes, including *beliefs*, *intentions* and *desires*.

2.2 Expressing attitudes and (in)sincerity

Broadly put, speech act theorists take a speech act to be insincere whenever there is a mismatch between *the attitude expressed by the utterance* and the *attitude possessed by the speaker* (Falkenberg 1988:93). Taken out of context, this definition is not very informative; in what follows, I will provide some theoretical background to flesh it out in a meaningful way.

It is a standard view in speech act theory that each illocutionary act expresses a distinctive propositional attitude (Searle 1969, Bach & Harnish 1979). The distinct attitude expressed by a given illocutionary act is part of what identifies it as opposed to others, and it is generally taken to define the point or purpose of the actions that we perform in uttering it. In this sense, we say that an assertion expresses a *belief*, that a promise expresses an *intention*, that asking someone for something expresses a *desire*. Philosophers and linguists have presented different taxonomies of illocutionary acts based, amongst other things, on the different psychological attitudes expressed by different (kinds of) illocutionary acts. Most authors would agree that the following characterisation is broadly⁵⁶ correct:

⁵⁶ It is debatable, for instance, whether a question or request always expresses a desire. For the purpose of the dissertation, however, we can leave this question aside. Independently of which

Attitudes expressed by specific illocutionary acts

B-EX: If S asserts that p, S expresses (EX) the belief (B) that p

D-EX: If S asks for p, S expresses a desire (D) for p

I-EX: If S promises that p, S expresses the intention (I) to do p

There are several ways to flesh out what is meant by ‘expressing’ a psychological state. I will assume that *expressing* an attitude does not entail having that attitude, so that you can insincerely express an attitude that you do not have (but cf. Davis 2003: 25, Green 2007:70-83). Following a hint by Davidson (1985:88, cf. Marušić 2012:13, Fallis 2013), one could say that for a speaker to express a psychological state is for the speaker to *represent himself* as being in that psychological state.

On an orthodox speech-act theoretic account of sincerity⁵⁷, the sincerity condition for uttering a given illocutionary act is that the speaker has the psychological attitude expressed by that act:

Sincere illocutionary acts

SIN: The performance of an illocutionary act $F(p)$ that expresses the psychological state $\Psi(p)$ is sincere IFF in uttering $F(p)$, S is in $\Psi(p)$

From the orthodox account of sincerity, a simple account of insincerity can be derived: a speaker is insincere whenever he is not in the psychological state that is expressed by the illocutionary act performed:

Insincere illocutionary acts

INS: The performance of an illocutionary act $F(p)$ that expresses the psychological state $\Psi(p)$ is insincere IFF in uttering $F(p)$, S is not in $\Psi(p)$ ⁵⁸

attitude a question expresses, what matters is that we can plug the correct characterisation of question into the general model that I am adopting.

⁵⁷ This view has been defended, under different guises, by Hare (1952:13,19–20,168–99), Searle (1969:60, 64–8), Wright (1992:14), Williams (1996:136), Moran (2005b), Green (2007:70-83).

⁵⁸ One might be wary of the chosen scope of the negation in INS, as we could that S is $\Psi\neg(p)$ rather than requiring that S is not in $\Psi(p)$ In section 5, I will address this kind of worry, and present an alternative version of this condition.

This gives us a provisional, simple formulation of the insincerity conditions for the three illocutionary acts we are considering as examples:

Insincerity conditions for specific illocutionary acts

BIC: S asserts that p insincerely only if S does not believe that p

DIC: S asks for p insincerely only if S does not desire p

IIC: S promises that p insincerely only if S does not intend to p

Importantly, an utterance can be insincere without being a lie. This is because an illocutionary act can be insincere without satisfying the assertion condition for lying. For instance, in asking you to do something that I do not desire, I may be insincere but I am not thereby lying. It should be thus kept in mind that this is a general account of insincerity, and that without a definition of assertion it does not alone provide us with sufficient conditions for defining lying.

2.3 Insincerity conditions for promising

Let us go back to the speech-act theoretic puzzle. The problem introduced by the UNFAITHFUL WIFE example is that there seem to be cases in which the speaker lies simply by lacking the intention to perform the promised action. In order to solve this puzzle, we need an alternative account of the insincerity conditions for lying. Since the proposed counterexample is about promising, to simplify this task I will start by developing an account for lying by promising. I will then show that this account generalises to every other illocutionary act that can be used to assert.

At this stage, on the table there are only two competing ways of accounting for the insincerity conditions for lying by promising. The *first* is the approach traditionally used to define lying: the *only-belief* account. This approach applies indifferently the same insincerity conditions to every illocutionary act, and counts an utterance as insincere whenever its content is believed to be false.

The *second* is the speech act theoretic account that I just introduced. According to this account, (i) a promise expresses an intention to Φ , and consequently (ii) the insincerity condition for promising is intending not to Φ . Assumption (i) can be traced back to Hume's view (*THN*: 517-19) that a promise always *expresses* (and communicates) an intention to

perform the promised act⁵⁹. Assumption (ii) is found in foundational works of speech act theory, like those of Austin (1962/1975:50, 135-6) and Searle (1965:243, 1969:60-2). I will refer to this account as the *only-intention* account. On this view, a promise is insincere only if, *at the time of the utterance*, the speaker does not intend to perform.

A *third* account of the insincerity conditions for promising can be derived by combining the previous two, in the light of the relation of entailment between asserting and promising that I outlined in III.3.2. As a reminder, the performance of an illocutionary act $F_1(p)$ entails the performance of another illocutionary act $F_2(p)$ iff in the context of the utterance it is not possible for S to perform $F_1(p)$ without performing $F_2(p)$ – so that if S performs $F_1(p)$, S also performs $F_2(p)$. This relation of illocutionary entailment occurs between assertions and promises: one cannot promise to Φ without also performing an assertion that one will Φ , so that every time one promises to Φ one also asserts that one will Φ . To recycle my previous example, my explicit promise (2) illocutionarily entails the assertion that (2*), since I simply cannot promise that I will feed the brontosaurus without thereby also asserting that I will feed the brontosaurus:

(2) I promise that (2*) [I will feed the brontosaurus]

How does this affect the insincerity conditions for promising? My conjecture is that if an assertion is always performed in addition to a promise, for the promise to be sincere the sincerity conditions for asserting need to be satisfied too. In other words, performing a sincere promise that Φ requires one both to intend Φ and to believe that one will Φ . This yields a novel account of the insincerity conditions for promising:

Entailed-Insincerity condition:

A promise is insincere if speaker intends not to Φ , or if the speaker believes that he will not Φ , or both

⁵⁹ This view is extremely influential in the philosophical literature on the nature of social obligations. Many authors (sometimes referred to as ‘information-interest’ theorist) take promising’s main function to be informing the promisee of what the promisor is going to do (Sidgwick 1981:442–44, Anscombe 1981:18, Rawls 1981:345, Fried 1981:16, Foot 2011:45). This view has been opposed by Owens (2008:747–51). His arguments seem successful in establishing that in promising to Φ one does not necessarily *communicate* an intention to Φ , but it is less clear that they demonstrate that in promising to Φ one does not necessarily *express* an intention to Φ , or that sincerely promising does not require intending to Φ .

More specifically, a promise to Φ is insincere *qua promise* if the speaker intends not to Φ , and insincere *qua assertion* if the speaker believes that he will not Φ . This view (be it correct or not) can clearly be generalised: whenever there is illocutionary entailment, and two illocutionary acts are performed, the sincerity conditions of both acts apply. To sum up, the three candidate insincerity conditions for lying by promising are:

Candidate insincerity conditions for promising:

(BIC): Belief insincerity condition: S believes that S will not Φ

(IIC): Intention insincerity condition: S intends not to Φ

(EIC): Entailed insincerity condition: BIC \vee IIC

Which account is preferable? The Moorean test for insincerity provides some linguistic data that *prima facie* favours the *entailed-insincerity* view over the other two. It is well known that assertions followed by the negation of their sincerity condition give rise to Moorean absurdities (Moore 1993:210): asserting “p and I don’t believe it” is incoherent in some distinctive way. One way of explaining this incoherence is that in uttering these sentences, the speaker performs a speech act and then blatantly violates one condition for its felicitous (in this case, sincere) performance, eventually failing to assert that p (Vanderveken 1980, Searle and Vanderveken 1985: 150-52). If both BIC and IIC are insincerity conditions for promising, it seems that they should both give rise to the same kind of unsuccessful incoherence (cf. Marušić 2012:14). As a matter of fact, sentences like (2- \neg B) and (2- \neg I) display this kind of absurdity:

(2- \neg B) I promise that I will pick you up at 6, but I don’t believe I will pick you up at 6 #

(2- \neg I) I promise that I will pick you up at 6, but I don’t intend to pick you up at 6 #

In both cases, the utterance strikes one as incoherent, and in both cases, it is difficult to imagine that the speaker will be taken to have promised to pick his interlocutor up at 6. This linguistic data supports EIC, and cannot be easily explained by BIC or IIC taken separately.

More importantly, only the EIC seems to make the right predictions in cases in belief and intention come apart, as in UNFAITHFUL WIFE. In UNFAITHFUL WIFE, Baba has a sincere belief that she will end up sticking to her promise (1), but an insincere intention to do whatever is in her power to break it.

(1) Do not worry Coco: I promise that I will not cheat on you tonight

While *the only-belief condition* BIC makes the incorrect prediction in this case, the *only-intention condition* IIC and *entailed insincerity condition* EIC correctly count (1) as a lie. The latter two conditions are thus preferable. But unlike IIC, EIC makes the correct predictions also when belief and intention diverge in the opposite way, *i.e.* when the speaker intends to do his best to stick to his promise, but believes he will end up violating it nonetheless. To see this, let us consider another example:

UNRELIABLE MECHANIC

Baba has broken her car, but she needs it to visit her family next week. For this reason, Baba has called Coco the mechanic to repair it. Coco the mechanic checks the car and tells Baba:

(3) Do not worry Baba: I promise that I will repair your car by next week

Coco intends to repair the car and he will attempt to do it, but he is almost certain that he won't manage to repair it in the end, because the damage is too serious.

In this example, Coco the mechanic promises to repair the car even if he knows that he will almost certainly fail to repair it: intuitively, (3) is a lie. In this case, the *only-intention condition* IIC is not able to track this intuition; whereas both the *entailed-insincerity condition* EIC and the *only-belief condition* BIC, correctly predicts that (3) is a lie.

Cases like UNFAITHFUL WIFE and UNRELIABLE MECHANIC support the conjecture that EIC is a better account of the insincerity conditions than BIC and IIC taken separately. Arguably, these cases are not straightforward, or prototypical, cases of lying. But this is also a prediction of the *entailed-insincerity* account: (3) is insincere *qua assertion*, but not insincere *qua promise*. Coco the mechanic intends to fulfil the promise, following one sincerity condition, but he believes he will almost surely fail, violating the other. By contrast, (1) is insincere *qua promise*, but not *qua assertion*. Baba intends to do her best to break the promise, violating one sincerity condition, but she believes that despite her efforts she will end up violating it, thereby following the other.

All in all, it seems that there are solid reasons to prefer the *entailed-insincerity* account of promising. So far, only Marušić (2013) has defended a similar view (Austin 1961:239 merely hints at this idea). However, Marušić also suggests that a promise that violates BIC

but not IIC might be better described as *irrational* rather than *insincere*. This is true if one endorses a ‘cognitivist view’, contending that it is irrational to intend to do something you believe that you will not do, so that only utterances satisfying both BIC and IIC (or neither of them) are rational, while utterances like (3) are *irrational* rather than *insincere*.

Marušić’s observation points out a possible problem for the proposed account: if a rational intention to Φ requires believing that one will Φ , then there is no need to require both IIC and BIC, as the satisfaction of the first entails the satisfaction of the second in every case in which the speaker is rational. I will not discuss this objection here; however, in Marsili (2016) I have argued that (as long as you take it as a live possibility that you will Φ – *i.e.* as long as you are not certain that you will not Φ) you can rationally intend to Φ and believe that you will very likely not Φ (or vice versa). On this *weak cognitivist* view, Marušić’s observation is not a worry: utterances like (3) are *insincere* rather than *irrational*, and EIC is preferable to the other accounts exactly because it successfully captures these peculiar forms of insincerity. In IV.3, I report empirical evidence showing that native English speakers judge that promises with contrasting intentions and beliefs are mendacious rather than irrational, and that they do not find the contrast between intention and belief involved in these cases to be problematic.

2.4 A general account of the insincerity conditions for lying

If my arguments are sound, it should be established at this point that EIC offers the best characterisation of the insincerity conditions for lying by promising. Pairing the EIC with the definition of lying developed in the previous chapter, we can obtain the following definition for lying by promising that p :

Definition of lying by promising

In successfully uttering a promise with content p , S lies to A about p iff:

1. *S thereby asserts that p*
2. *Either S believes that not p , or S does not intend to p , or both*

Since promises by default entail assertions, condition (1) obtains by default: the informative bit is condition (2), that specifies under which conditions a promise is a lie. This definition can then be extended to other illocutionary acts that entail an assertion. We know from the general account of insincerity developed in the previous section that an illocutionary act

$F(p)$ that expresses the psychological state $\Psi(p)$ is insincere IFF in uttering $F(p)$, S is not in $\Psi(p)$. To generalise the definition of lying by promising to illocutionary acts other than assertion, we simply need to require that, when an illocutionary act other than assertion is performed, either this insincerity condition is satisfied, or the insincerity condition for assertion, or both. In other words:

Definition of lying by performing an illocutionary act

In successfully uttering an illocutionary act with content p that expresses an attitude $\Psi(p)$, S lies to A about p iff:

1. *S thereby asserts that p , i.e.:*
 - a) *S expresses p*
 - b) *S presents p as an actual state of affairs*
 - c) *S takes responsibility for p being an actual state of affairs*
2. *Either S believes that not p , or S is not in $\Psi(p)$, or both*

This gives us a more accurate account of the conditions under which an illocutionary act other than assertion can count as lying. This definition is more accurate than the one developed in chapter III, as it generalises the insincerity conditions to attitudes other than beliefs. But does this definition really reflect our ordinary intuitions about lying? In what follows, I present evidence for a positive response to this question.

3. INSINCERE PROMISES: AN EXPERIMENTAL STUDY

3.1 Testing folk intuitions about lying

In the philosophical literature, it is generally agreed that a good definition of lying should track the ordinary usage of the term (Carson 2006:285; Fallis 2009:32): most debaters are after a characterisation of lying that is in line with the linguistic practice of competent speakers. A good account of lying should predict which usages of the term are correct and incorrect according to competent speakers. A corollary of this way of thinking is that if an account of lying makes predictions that are inconsistent with ordinary people's intuitions, that account fails to meet one important *desideratum* of a theory of lying. With this in mind, philosophers and linguists have started to accumulate data about ordinary speakers' intuitions about the correct usage of the terms. These studies have the potential to give us insight into what lying is, or at the very least into what lying is perceived to be within a community of speakers⁶⁰.

Numerous and diverse empirical studies, stemming from different theoretical backgrounds and motivated by different explanatory aims, have attempted to explore folk intuitions about lying⁶¹. Among the ones explicitly investigating the intuitions of competent speakers about the concept of lying, the most important strand comes from the framework of prototype semantics. Following the lead of Coleman & Kay (1981), these studies attempt to identify the features of a prototypical lie, and to outline the differences of these prototypes across different cultures (Sweetster 1987, Cole 1996, Hardin 2010, Rong, Chunmei & Lin 2013). The present study addresses similar worries, but in a slightly different framework, namely that of experimental philosophy. Here, the aim of the analysis is to identify the necessary and sufficient conditions for an utterance to be a lie, rather than the prototypical features that make up the concept. Only a few studies on lying have been conducted within this framework so far, some attempting to test if the intention to deceive is necessary for lying (Arico & Fallis 2013, Meibauer 2016, Rutschmann, R., & Wiegmann, A. 2017), and others if actual falsity is (Turri & Turri 2015, Wiegmann et al. 2016). The

⁶⁰ For a more detailed defence of the importance of tracking ordinary intuitions for a definition of lying, see Fallis (2009) and Arico & Fallis (2012:794-7)

⁶¹ Here I am only considering studies on competent speakers. For a broader a review, including studies in developmental psychology, see Hardin (forth.).

present experimental study will instead try to establish which conditions are necessary for lying by promising.

3.2 Aim of the study

This experimental study aims to test the theories developed in section 2 against the intuitions of native English speakers. This means that it will attempt to determine if ordinary people rate illocutionary acts other than assertions as lies, and which *insincerity conditions* have to be satisfied for them to do so. More specifically, this study will be concerned with one speech act in particular, namely promises. The main reason is that promises clearly display all of the speech-act theoretic puzzles introduced in chapter III: they can be performed by means of an explicit performative, they are not assertions, and their insincerity conditions are sensitive to attitudes other than belief (namely intentions). As a reminder, in the previous section I have introduced three candidate insincerity conditions for promises: BIC IIC or EIC:

Candidate insincerity conditions for promising

IIC: the speaker does not intend to Φ

BIC: the speaker does not believe that he will Φ

EIC: either the speaker does not believe that he will Φ , or does not intend to Φ , or both

Testing that illocutionary acts other than assertions are judged to be lies is relatively simple: it is sufficient to create a story in which it seems that a character lies by promising, and ask participants whether the character has lied. If participants classify promises as lies, we have evidence that not only assertions are classified as lies. Testing which insincerity condition is more accurate is a slightly more complex matter. Here we need different stories in which different combinations of insincerity conditions are violated, and for each story test whether the participants believe that the speaker has lied. Given our candidate insincerity conditions, we will need to consider three scenarios in which a character promises something insincerely. In the *straightforward scenario*, the character's (S) utterance satisfies both BIC and IIC. In the *no-intention* scenario, it satisfies IIC but not BIC. In the *no-belief* scenario, it satisfies BIC but not IIC. I will refer to the latter two cases as the *crucial* conditions, as opposed to the *control* (straightforward) conditions.

- *Straightforward scenario*: BIC & IIC [CONTROL]
- *No-intention scenario*: IIC & ¬BIC [CRUCIAL]
- *No-belief scenario*: BIC & ¬IIC [CRUCIAL]

3.3 The predictions of existing theories

In section 2 I have mentioned five approaches to lying by explicit promising. Each account gives different predictions about which of the three scenarios will be rated as a lie.

1. According to the *only-assertion paradigm* (\emptyset), lying requires a direct assertion. Its prediction is that since promises are not direct assertions, respondents will claim that in no scenario is the character lying. Interestingly, if this view is correct, the first two speech-act theoretic puzzles are not really a worry for defining lying: since only direct assertions can count as lies, there is no issue of determining whether explicit performatives can count as lies, and under which conditions.
2. The *only-belief paradigm* (BIC) rigidly assumes that you lie only if you *believe* that the propositional content of your speech act is false. This view expects positive responses in the *straightforward* and in the *no-belief* condition, but negative responses in the *no-intention* condition. If this view is correct, the third speech act theoretic puzzle about insincerity is not really a worry for defining lying: since only beliefs are relevant to determine whether an utterance is a lie, there is no issue of extending the insincerity conditions to other attitudes.
3. The *only-intention paradigm* (IIC) maintains that a promise is insincere iff the speaker does not intend to fulfil his promise. Applied to lying, this view predicts that a promise is a lie iff the speaker does not intend to fulfil it. The *straightforward* and *no-intention* cases should then be rated as lies, but not the *no-belief* case.
4. According to a *cognitivist interpretation* (BIC & IIC), the *no-intention* and *no-belief* should be described as cases of irrational thinking rather than lying, so that only the *straightforward* case should be rated as a lie.
5. According to the *entailed-assertion paradigm* (BIC \vee IIC), a promise is a lie either if the speaker does not intend to fulfil his obligation, or if he believes that he will fail to fulfil it, or both. The account also predicts that all scenarios will be rated as lies, but expects the *straightforward* one to receive slightly higher ratings than the *crucial* ones.

| | STRAIGHTFORWARD | NO INTENTION | NO BEL |
|--------------------|-----------------|--------------|----------|
| ONLY-ASSERTION | NO | NO | NO |
| ONLY-BELIEF | YES | NO | YES |
| ONLY-INTENTION | YES | YES | NO |
| COGNITIVIST | YES | NO (IRR) | NO (IRR) |
| ENTAILED-ASSERTION | YES | YES (-) | YES (-) |

Table 1: the predictions of the five different accounts of lying by promising

3.4 Experiment 1

3.4.1 Method

Participants: Participants were recruited using Amazon Mechanical Turk and tested using Qualtrics. They were compensated \$0.2 for taking the survey. Repeated participation was prevented. Overall, 166 U.S. residents were tested (85 females; mean age (SD) = 36.6 years (12.9); range: 18–72; 100% reporting English as a native language). To prevent participants from taking the test negligently, a minimum response time (25 seconds) and a control question were set. Data from three participants who failed to meet these conditions were excluded, but including them would not affect the results.

Design: Each participant was randomly assigned to one of four conditions. Each condition features Coco and Baba, and in each condition Coco promises something to Baba. The first two conditions [(1) *straightforward lie*, (2) *no intention*] belong to the ‘drink’ story, and the second two conditions [(3) *straightforward lie*, (4) *no belief*] belong to the ‘repair’ story. In the ‘drink’ story, Coco promises not to drink; in the ‘repair’ story, Coco promises to repair Baba's car. For each pair, in the *straightforward* condition Coco lacks both intention and belief, and in the crucial case he lacks one attitude (intention in 2, belief in 4) but not the other.

Having been assigned to one of the conditions and having read the relevant story, the participants were posed two questions, always in the same order. The first asked whether Coco told a lie (“Did Coco tell a lie”? Y/N). The second allowed participants to report whether they felt uneasy in answering the preceding question (“Did you find it easy to make a decision”? Y/N); those answering “no” were solicited to motivate their uneasiness via a simple feedback form. The second question was especially designed to rule out the

possibility that participants would have preferred not to give a dichotomic yes-no response, but also to collect qualitative data about the strength of the participants' intuitions.

Some peculiarities of the design are due to consistency constraints on the rationality of “intending without believing”, and vice versa (cf. Appendix II). The first peculiarity is that in all conditions Coco has a partial rather than outright belief in whether he will fulfil his promise. This is due to the adoption of *weak cognitivism* in this research, according to which intending to Φ entails not being certain that you will fail to Φ . To grant uniformity between all conditions, Coco has a partial belief both in the crucial cases (where it could not be otherwise) and in the straightforward ones. Similar consistency constraints (also discussed in Appendix II) motivated one asymmetry in the experimental design, *i.e.* the fact that the no-intention and no-belief cases were not tested within the same story. The reason is that an uncontroversial *no-intention* case demands a promise about *refraining from acting*, while an uncontroversial *no-belief* case demands a promise about *actively performing one action*.

3.4.2 Results and discussion

Virtually every respondent rated the *straightforward cases* (in which Coco lacks both belief and intention to perform) as lies: 95% of the participants claimed that Coco lied in the *drink-straightforward* condition (38 of 40) and 95% in the *repair-straightforward* condition (39 of 41). All except one participant (in the drink scenario) declared that the question was easy to answer. The results for the straightforward cases support the view that it is possible to lie by explicit promising, refuting the *only assertion* hypothesis. They also back the stronger claim that insincere promises can be regarded as *prototypical* cases of lying; and that, more generally, a prototypical lie can be performed by uttering a speech act other than assertion.

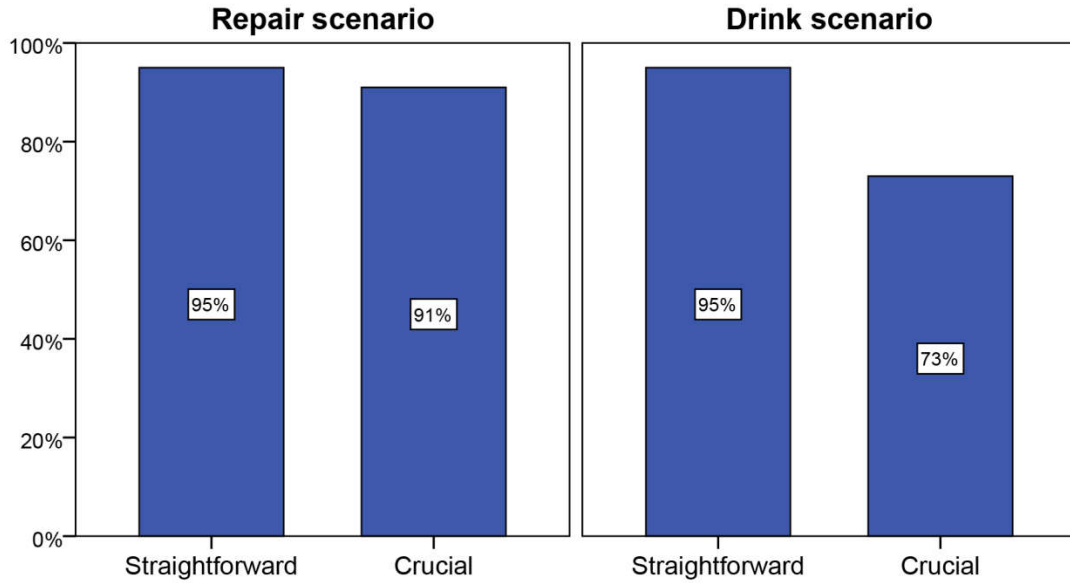


Figure 1: Percentage of respondents rating the protagonist's utterance as a lie in each condition.

In the *no-intention* condition, 90% of the participants (40 of 44) rated the promise as a lie. This refutes the predictions of the *only-belief* and of the *cognitivist* accounts, as respondents classed the promise as a lie even if Coco sincerely believes that he will (almost certainly) fulfil his promise. The difference between this condition and the corresponding *drink-straightforward* case was not significant, $\chi^2(1, N = 84) = .53, p = .467$; nonetheless, as much as 14% of the participants (6 of 44) declared that the question was not easy to answer, which is significantly more than in the *straightforward* conditions (Fisher's Exact Test, $N = 125, p = .008$, two tails). This shows that intuitions in this condition are not as strong as in the *straightforward* cases, and suggests that the *no-intention* condition is regarded by some participants as a non-paradigmatic instance of lying.

The *no-belief* condition was rated as a lie by 73% of the participants (30 of 41). A binomial test showed that this result is significantly different to a chance distribution of the responses ($p = .004$, two tails); in other words, participants were more likely to say that the promise was a lie than the opposite (OR = 2.73), which logically entails that the *only-intention* account (according to which this case does not qualify as lying) can also be rejected. In this condition, 15% of the participants (6 of 41) declared that the question was difficult to answer: Fisher's Exact Test revealed that this was significantly different from the *straightforward* cases ($p = .006$, two tails, OR = 13.71), suggesting that at least some participants did not see this as a paradigmatic case of lying. That the case might not be seen as prototypical is also confirmed by the lower percentage of ratings of the promise as a lie:

this case differs significantly from all the rest of the cases jointly, $\chi^2(1, N=166) = 12.71, p = .001$, as well as from the no-intention case separately, $\chi^2(1, N=85) = 4.6, p = .032$.

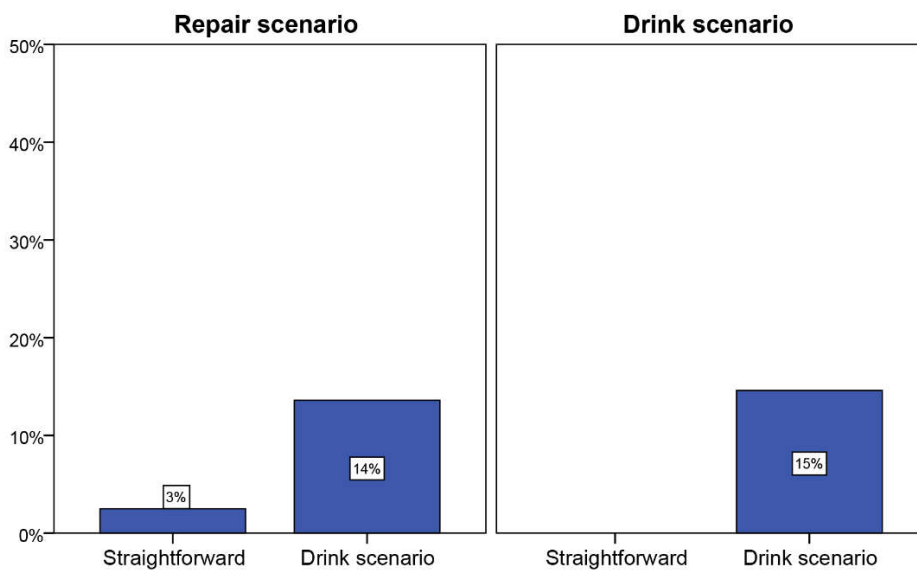


Figure 2: Percentage of respondents rating the question as difficult to answer (scale from 0% to 50%).

Interestingly, none of the feedback forms filled in by the participants who found it difficult to answer the question mentions the contrast between intention and belief as problematic, suggesting that the *cognitivist account* of intentions does not reflect the participants' intuition. Logistic regression was used to test whether information about gender or age could be used to predict the responses, revealing that neither gender ($p = .592$) nor age ($p = .808$) was a significant predictor.

The results provide strong support for the definition of lying by promising elaborated in part A. However, one can still worry that the dichotomous yes/no design led some participants to polarize their responses in an unnatural way (cf. Xu, Lu, Fuo & Lee 2009:318, Turri & Turri 2015:163). In section 3 we have seen that beliefs, and hence insincerity, can come in degrees, and that contrasting parameters can influence our evaluation of a statement as insincere. Perhaps participants see lying as a graded or vague concept (cf. IV.4 and Bazzanella 2009) and would have preferred to give a graded answer. In fact, 14% of the participants in the crucial cases (12 of 85) declared that the question was difficult to answer. Maybe they gave a positive answer to sanction that the speaker misbehaved in some important way – but they would not have classed the utterance as a lie if an intermediate alternative was given. If this line of reasoning is correct, and the probes forced the participants to polarise their responses, the proposed interpretation might be

mistaken. To test this hypothesis, the test was rerun with a slightly modified design that allowed for graded rather than dichotomous responses.

3.5 Experiment 2

3.5.1 Method

Participants: Participants were 104 adults (57 females; mean age (SD) = 37.1 years (11.8); range: 19-64; 100% reporting English as a native language, recruited and tested online as in Experiment 1). One participant failing to meet the minimum response time was excluded, but including him would have not affected the results.

Design: The design was the same as Experiment 1, but instead of using a dichotomous yes/no measure, participants were asked to rate their agreement with the statement “Coco told Baba a lie” on a 7-point Likert scale (1: “definitely agree”, 4: “neither agree nor disagree”, 7: “definitely disagree”). If people had the intuition that the crucial conditions do not fully qualify as cases of lying, their ratings should concentrate around the midpoint (or below); if instead they shared the intuition that Coco is actually lying, the ratings should instead concentrate on the ‘lie’ end of the scale (5 or above).

3.5.2 Results and discussion

In the *no-intention condition*, participants overwhelmingly agreed that Coco did lie: 84% (21 out of 25) of them rated it as “5” or above, and only 16% (4 participants) as “4” or below (mean = 6.0; mode = “7”). Similarly, in the *no-belief condition*, participants overwhelmingly agreed that Coco did lie, 88% (22 out of 25) of them rating it as “5” or above and the remaining 12% (3 participants) rating it below “4” (mean = 5.8; mode = “7”). Like in the first experiment, the mean score in the ‘repair belief’ condition was slightly lower than in the ‘drink intention’ condition, but the difference (unlike in Experiment 1) was not significant, $t(48) = 0.64$, $p = .523$. A comparison of the crucial and the straightforward cases revealed a significant difference between the two ‘repair’ scenarios, $t(34.3) = 2.52$, $p = .017$ (equal variances not assumed), but no significant difference between the two ‘drink’ scenarios, $t(49) = 0.58$, $p = .562$ ⁶². Overall, these results are well in accordance with those

⁶² Note, however, that this comparison may have been somewhat distorted by the fact that the control drink scenario got lower scores than expected from a straightforward case.

obtained with the dichotomous design, strengthening the case for the *entailed-assertion* account.

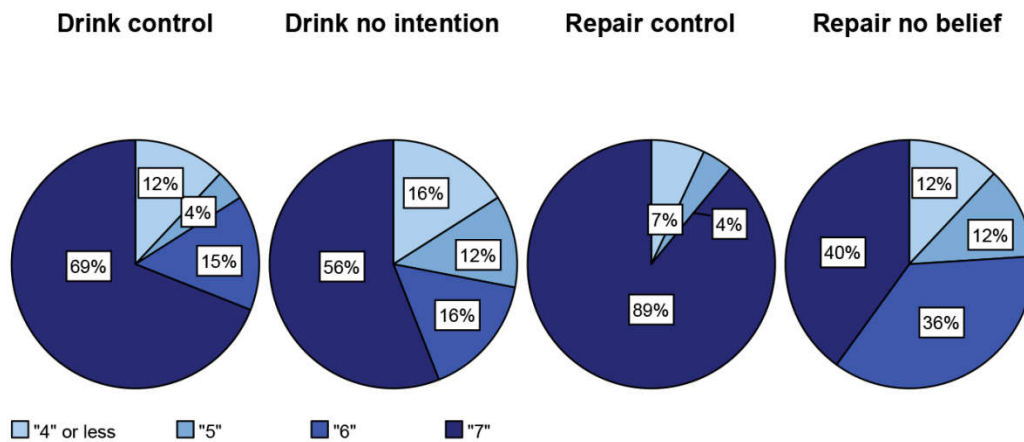


Figure 3: Pie charts showing the participants' ratings in each of the four scenarios.

A last worry to address is that participants might have agreed that Coco lied only because they were not allowed to categorise his statement in some alternative way that they found more adequate. Perhaps most had the intuition that in the crucial conditions Coco was deceptive, or insincere, and were led to describe him as lying only because no other category of assessment was offered. They agreed that Coco was lying because it was the only available option to express that he misbehaved, but they would have denied it if some alternative category more adequate to describe the situation, like being deceptive, was also available. To address this worry, a second study was conducted that provided participants with an opportunity to describe the speaker along two different categories: as being deceptive and as being lying⁶³.

3.6 Experiment 3

3.6.1 Method

⁶³ What if the participants preferred to describe the utterance as insincere rather than deceptive? There is a reason why “deceptive” was preferred to “insincere”. On all plausible understandings of these terms, being insincere entails being deceptive, while the opposite is not true. The “deception” option is thus preferable, as it allows all participants to acknowledge that the protagonist misbehaved also if they think that “insincere” is a more accurate description (agreeing that the protagonist is insincere entails agreeing that he is deceptive).

Participants: Fifty-five new participants were tested (31 females, mean age (SD) = 34.6 years (12.5); range: 19–66; 100% reporting English as a native language, recruited and tested online as in the other experiments). Data from six participants failing to meet the minimum response time and/or the control question was excluded, but including them would have not affected the results.

Design: Participants were randomly assigned to one of the two crucial conditions (*no-belief*, *no-intention*) from Experiment 1. They were then asked to answer two questions that appeared (in randomized order) on the same screen:

- Did Coco say something deceptive? [Y/N]
- Did Coco tell Baba a lie? [Y/N]

Participants were then asked to answer the same control and demographic questions as in Experiment 1.

In both of the vignettes used in the experiment, there is no question that Coco's statement is deceptive: in the *no-belief* condition he pretends to have a belief he does not have; in the *no-intention* one, he pretends to have an intention that he does not have. Participants who had the intuition that Coco was *not lying* thus had the opportunity to deny that Coco is lying while being able to describe Coco as misbehaving, namely as being deceptive.

Given that in both vignettes it is uncontroversial that Coco's statement is deceptive, it is only the second question (about lying) that is of interest here. If the scores obtained in this experiment are significantly lower than those obtained in the same scenarios of Experiment 1, then the proposed interpretation of the results might be unwarranted. By contrast, if participants continue to describe Coco as lying, there is even stronger experimental evidence in favour of the proposed view.

3.6.2. Results

The overwhelming majority of participants marked both the *no-intention* case (92%, 24 out of 26) and the *no-belief* case (90%, 26 out of 29) as a lie. The scores obtained in the crucial conditions in Experiment 3 are even higher than those obtained in Experiment 1, but neither of the across-experiment differences was significant (*no-intention*: $\chi^2(1, N = 70) = 2.89, p = .089$; *no-belief*: $\chi^2(1, N = 70) = .04, p = .84$). These results are consistent with the previous ones, and strengthen the case for the *entailed-assertion* view. Somewhat

surprisingly, the difference between the *no-belief* and the *no-intention* condition found in the previous experiments has disappeared, $\chi^2(1, N = 55) = .12, p = .73$.

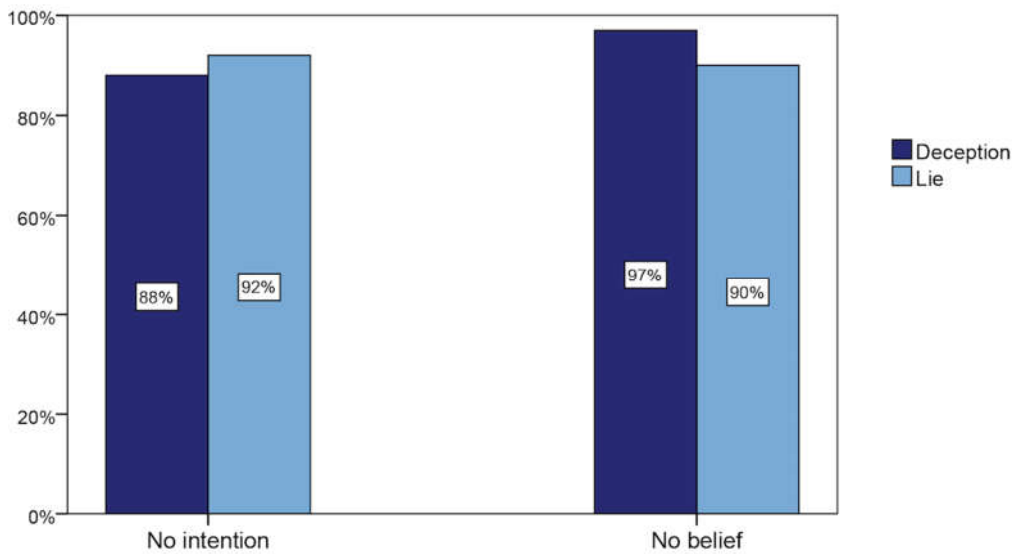


Figure 3: Percentage of respondents rating the protagonist's utterance as a lie and as deceptive.

3.7 General discussion

The results of all experiments strongly support the *entailed-assertion* account. They consistently show that it offers the best predictions of people's intuitions about lying by promising. A promise is a lie iff the speaker lacks belief or intention to fulfil the promise, or both. Furthermore, these results undermine all alternative views: the *only-assertion* view that only assertions can count as lies, the *only-belief* view that believing that what you say is false is necessary for lying, and the *only-intention* view that a promise is a lie only if the speaker does not intend to perform. In doing so, they also confirm the significance of the speech-act theoretic puzzles about lying, as these puzzles are directed at criticising both the *only-assertion* and *only-belief* view about lying.

In Experiment 1 and 2, the crucial conditions obtained lower scores than the control conditions (Experiment 3 had no control conditions), and in Experiment 1 around one in six participants declared that they found it difficult to decide whether to classify them as lies. In line with the prediction of the *entailed-assertion* view, this suggests that the consensus is less pronounced when only one of the two sincerity conditions for promising is violated, and that a portion of the population sees them as non-paradigmatic instances of lying.

3.8. More on the results

3.8.1 Intentions vs Beliefs

The crucial conditions got different results in Experiment 1: the *no-intention* condition got significantly higher rates than the *no-belief* one. A slighter difference between them was also found in Experiment 2. This kind of difference, however, was not found in Experiment 3, or in the second question of Experiment 1 (about the participants' uneasiness to define Coco's utterance as a lie). Perhaps the difference is due to random fluctuations in the subjects' intuitions, and does not require an explanation. But if an explanation must be given, it can be given in terms of *moral judgements*.

Several authors have suggested that lying is a morally loaded term (Bok 1978:14, Williams 1985:140), and some of them even contended that white lies are not lies (Margolis 1962, Donagan 1977:89, Grotius, RWP:1212-8). A plausible view is that moral judgements might affect whether one finds a particular case of lying more or less prototypical. Experimental studies have also shown that judgements about intentions (Knobe, 2003) and causation (Alicke 2014) can be influenced by judgements about culpability (*i.e.* moral judgements). If moral judgements can affect folk intuitions about whether a particular case is a lie, a moral asymmetry between the scenarios might have influenced the results. In fact, there is such an asymmetry between the *no-intention* and the *no-belief* cases. In the first case, Coco is fully responsible for not fulfilling the promise: it is in his power to fulfil it, but he willingly decides to infringe it. By contrast, in the *no-belief* case Coco intends to do what is in his power to fulfil the promise, but it is not fully in his power to do so. In other words, in the first case he is fully responsible for the infraction, while in the second he is only responsible for having set the stakes too high.

While the *no-intention* case is by definition in 'bad faith', the *no-belief* case is by definition in 'good faith'. This could explain the slight asymmetry in the results - an asymmetry also expected in any replication of the test, given that it is built into the difference between violating the 'belief sincerity condition' and violating the 'intention sincerity condition'.

3.8.2 The intention to deceive condition

Proponents of the *intention to deceive condition* (IDC) for lying (cf. chapter II) might be worried that these results have been influenced by the fact that it has been left unspecified whether Coco intends to deceive Baba. For instance, the *no-belief* case might have received lower results in Experiment 1 because it is not clear whether Coco intends to deceive Baba.

In this section, I will show not only that this worry is unfounded, but that the results undermine the very idea that ordinary speakers take the IDC to be necessary for lying.

On a standard interpretation (the “rigid” intention to deceive condition, or IDC2, cf. II.2.2), the relevant intention to deceive has to be about the content of the statement: if the content of the statement is p , the speaker has to intend to make the hearer believe that p . On a weaker version of the IDC (the “believed sincerity condition”, or BSC, cf. II.2.3), the relevant intention is just to make the hearer believe that *the speaker believes* that p . In other words, where p is the *believed-false* content of the speaker’s statement, the two most influential versions of IDC are:

IDC2: S intends A to believe: (p)

BSC: S intends A to believe: (S believes that p)

In the *no-belief* condition, the content of the promise is that Coco will repair the car, a proposition that Coco believes to be very likely false. Now, unless Coco intends his promise not to be accepted, or not to be acted upon (*i.e.* if Coco’s promise is a normal promise), Coco clearly intends Baba to believe that he will repair the car, so the participants have no reason to believe that IDC does not obtain.

In the *no-intention* condition, by contrast, the relevant content is that Coco will not drink. But here Coco believes the proposition to be probably true, so that neither IDC2 nor BSC can obtain, and the participants cannot think that they obtain. Two consequences can be drawn from this observation. The first one is that intuitions about the intention to deceive condition did not alter the result of the experiment: if they had some weight, they would have favoured the *no-belief* case over the *no-intention* one; instead, it was the latter that obtained significantly higher results. The second is that the high results from the *no-intention* case (90% lies, showing no significant difference from the straightforward cases) strongly suggest that neither IDC2 nor BSC is perceived as a necessary condition for lying. Here is a reply. Even if neither IDC1 nor BSC are satisfied in the *no-intention* case, Coco is still aiming to deceive, since he clearly intends Baba to believe that he intends not to drink. The problem with this response is that it relies on a problematic definition of ‘intention to deceive’ (paralleling the “broad” intention to deceive condition, or IDC1, from II.2.1) according to which there are no constraints on what the deception is about:

IDC1: S intends to deceive A

The problem with the IDC1 is that we have already seen that it is untenable, because it counts any deceptive believed-false statement as a lie, even if deception has nothing to do with the content of the statement. It seems that the IDC should instead only capture deceptive intents that are somehow related to what is said by the liar. As a matter of fact, there are known counterexamples to IDC1, as the THEATRICAL POSE counterexample by Fallis (2010:6) (cf. II.2.1).

To be able to defend the claim that Coco is lying in the *no-belief* case, the proponent of the IDC has to provide a different version of the IDC – perhaps one that is sensitive to the different attitudes expressed by a speech act, such as IDC4:

IDC4: S intends A to believe that (S Ψ (p)),

where Ψ is the propositional attitude expressed by the illocutionary act with content *p* that S has performed.

To conclude, the experiment represents a challenge to the existing versions of the IDC, and suggests a further challenge for their proponents: that of constraining the content of attempted deception in a way that generalises across different illocutionary acts.

3.8.3 Falsity condition

In the literature on lying, virtually every author accepts a ‘*subjective*’ account of lying, according to which asserting an *objectively false* proposition is not necessary for lying, as long as the speaker *believes* that proposition to be false (cf. 1.1.2). However, a few authors (Grotius RWP:1209, Benton forth) have suggested that falsity of the statement, in addition to belief in its falsity, is required for lying – call this the *objective* view. Recently, Turri & Turri (2015) claimed to have found experimental evidence that most laypeople endorse the objective view. However, their study is far from convincing, and has recently been dismissed by Wiegmann et al. (2016). An interesting aspect of the present study is that it puts further pressure on the objective view of lying, and on Turri & Turri’s claim that such view reflects laypeople’s intuitions.

How does the present study relate to this debate? A first suggestion is that if participants regarded falsity as necessary for lying, they would not have rated a promise as a lie unless the promise was actually infringed in the story. However, in all conditions of all experiments participants agreed that the protagonist lied, even if it is always left unspecified whether the

promise was fulfilled or not, *i.e.* if it was objectively false⁶⁴. This seems incompatible with the fact that all respondents in the straightforward conditions, and the majority in the others, indicated that Coco lied. Moreover, in the straightforward cases, 99% of the participants reported that they found it easy to respond. How could this be, if they did not know whether the falsity condition obtained?

One easy response is that in all four cases the respondents predicted that, given the information available, the promise would eventually be infringed in the story: they took it that it was ‘implicit’ that the falsity condition would obtain. This is clearly plausible for the *straightforward* conditions: since Coco intends not to do what he promised, and believes that very probably he will succeed in not doing it, the falsity condition will almost surely be met. A similar inference is plausible in the *no-belief* condition: even if Coco intends to repair the car, he believes that very likely he will not succeed, and this clearly suggests that the car will not be repaired.

The real problem for proponents of the falsity condition is the *no-intention* case: here, Coco intends to drink against his promise, but he believes that he will very likely fail to do so, because he will not be able to. The information provided in the scenario cannot license the inference to the conclusion that he will drink; as a matter of fact, it only licenses the opposite inference. In other words, not only it is not specified if the falsity condition obtains, but the scenario clearly suggests that it will not obtain. In all experiments, the *no-intention* condition was consistently rated as a lie (obtaining even higher scores than the other crucial condition), and in no experiment its results were significantly different to the straightforward scenarios: this strongly suggests that participants did not take falsity of the promise to be necessary for lying.

⁶⁴ What do we mean exactly by saying that such a promise *about a future state of affairs* can be false? No straightforward response can be offered in our case, because promises are about future contingents (at the moment in which a promise is uttered, it is still indeterminate whether the promisor will fulfil it: in some possible futures he does, in others he does not) and semanticists disagree about how to determine the truth-conditions of statements about future contingents. The fact that we are considering a promise rather than an assertion further complicates the issue. However, for our purposes it is sufficient to point out that no plausible account is able to predict that the falsity condition is met in the *no-intention* case. Having noted that, it is worth offering a sketch of how a plausible characterisation of the falsity condition for promising could look like. The following is broadly inspired on Belnap’s (2000) account of the truth conditions for promising that Φ :

The falsity condition for promising that Φ is met at the moment of the utterance m_U and at the relevant moment m_R ($m_R \geq m_U$) iff it is a settled matter at m_R that Φ was false at m_U .

I am grateful to Francis Cockadoodledoo for having helped me to develop this point.

The experimental data collected thus strengthens the case for the subjective account. As already shown by Wiegmann et al., people think that lying does not require falsity after all⁶⁵.

3.9 Conclusions

In the previous chapter, I have developed a speech-act theoretic framework to outline the necessary and sufficient conditions for lying by promising, and sketched a way to extend this account to other illocutionary acts performed by uttering a sentence containing an explicit performative. Assuming that an ‘assertion-based’ definition of lying is broadly correct, the proposed account takes a promise to be a lie under the following conditions:

Definition of lying by promising

In successfully uttering a promise with content p , S lies to A about p iff:

1. *S thereby asserts that p*
2. *Either S believes that not p , or S does not intend to p , or both*

More generally,

Speech-act theoretic definition of lying

In successfully uttering an illocutionary act with content p that expresses an attitude $\Psi(p)$, S lies to A about p iff:

1. *S thereby asserts that p , i.e.:*
 - a) *S expresses p*
 - b) *S presents p as an actual state of affairs*
 - c) *S takes responsibility for p being an actual state of affairs*
2. *Either S believes that not p , or S is not in $\Psi(p)$, or both*

In introducing my speech act-theoretic definition of lying, I have put into question several ideas that are found in the literature: the view that only direct assertions can be lies; the view that the propositional content of a lie has to be believed to be false; the view a promise is insincere only if the speaker does not intend to perform; the view that falsity is required for lying. My experimental study has shown that, at least with respect to promises, the proposed

⁶⁵ It should be stressed, however, that the experimental design of this study was not conceived specifically to test the falsity condition hypothesis. Even if the data collected supports the *objective* view, it provides weaker support than studies explicitly conceived to test this hypothesis, *i.e.* studies explicitly stated that the falsity condition did not obtain.

view (but none of the alternative ones) gives predictions that are consistent with folk intuitions about lying.

Overall, it seems that my proposed definition offers a plausible account of the conditions under which the performance of an insincere speech act counts as a lying, that reflects ordinary people's judgements about whether a given utterance is a lie.

Appendix I - the scenarios

Drink story - Refraining from action

Baba and Coco are married. Coco is going to a party, but Baba does not like Coco to drink. So Baba asks him: "Will you be drinking alcohol at the party?". Coco replies: "Do not worry Baba: I promise that I will not drink alcohol at the party."

Scenario 1 - Straightforward:

In fact, Coco intends to drink alcohol at the party, and he is almost certain that he will find something to drink there.

Scenario 2 - No intention (to refrain as promised):

Coco actually intends to drink alcohol at the party and he will attempt to, but he is almost certain that he won't succeed, since he believes that the hosts do not offer alcoholic drinks at their parties.

Repair story - Positive action

Baba has broken her car, but she needs it to visit her family next week. For this reason, Baba has called Coco the mechanic to repair it. Coco the mechanic checks the car and tells Baba: "Do not worry Baba: I promise that I will repair your car by next week."

Scenario 3 - Straightforward:

Coco has no intention whatsoever to repair the car, and he is almost certain that he will not repair it.

Scenario 4 - No belief:

Coco intends to repair the car and he will attempt to do it, but he is almost certain that he won't manage to repair it in the end, because the damage is too serious.

4. GRADEDNESS

The previous sections have extended the insincerity conditions for lying to attitudes other than assertions. This is an important advance in the definition of lying, but it still overlooks one often-undervalued feature of insincerity: the fact that it comes in degrees. As Montaigne nicely stated, while truth is unique, “the opposite of truth has many shapes, and an indefinite field” (Montaigne *E*: 1.IX). There is a whole grey area of ‘half-sincere’ utterances that are difficult to classify and, quite importantly, it is in this grey zone that liars strive.

To shed some light on this obscure area, this section will consider cases involving partial insincerity; for example, statements that are not fully believed to be false, but that are nevertheless not believed to be true. Are these statements lies? And how much confidence in their falsity is required for them to count as lies? We will discuss such questions, and explore the thin, elusive line that distinguishes a sincere assertion from an insincere one. This will be a hard challenge, and indeed for a theory of lying “the more difficult task [is] that of drawing lines” (Bok 1989:49).

The standard, simplistic account of the insincerity condition for lying is that an utterance is a lie only if the speaker believes it to be false. However, the expression “believe to be false” is not really helpful to deal with intermediate cases, as it does not specify what degree of confidence in the falsity of p counts as believing it be false. In what follows, I will develop a version of the insincerity condition that be able to classify statements that are neither fully believed to be false, nor fully believed to be true. For most of my discussion, I will pretend to adopt the *only belief* view: this will allow me to provisionally set aside the intricacies that arise when we consider attitudes other than belief. In the last section, I will extend the analysis to other attitudes, to provide a final, complete account of the insincerity conditions for lying.

4.1 The dichotomic view and the traditional insincerity condition

Both deceptionist and non-deceptionist typically agree on the following formulation of the insincerity condition for lying (call it the ‘traditional insincerity condition’, or TIC):

(TIC) S believes that p is false

Scholars endorsing this condition tend to take for granted that a statement is *sincere* when the speaker believes it to be *true*, and *insincere* when the speaker believes it to be *false*⁶⁶, and that a more fine-grained analysis would be unnecessarily intricate (Saul 2012:5, fn10). From this perspective (call it the *dichotomic view*), the definition of lying correctly rules out only statements that are believed to be true.

THE DICHOTOMIC VIEW

A statement is *sincere* when the speaker believes it to be *true*, and
insincere when the speaker believes it to be *false*, *tertium non datur*

It is not obvious, however, that the dichotomic view is correct, nor that TIC offers a satisfying characterisation of the insincerity condition for lying (Mahon 2015:1.5). There is good ground to suspect that the dichotomic view is not an adequate assumption defining lying, because a number of intermediate credal states⁶⁷ exist between believing *p* to be true and believing *p* to be false.

First, it is possible for a speaker to believe that a statement is only partly false (rather than utterly false): in this case, the speaker believes that *p* has a ‘graded truth value’. Second, it is possible for a speaker not to be certain; in other words, to have a graded degree of confidence (rather than a flat-out belief) in the falsity of a statement: intermediate beliefs of this kind are called ‘graded beliefs’. The difference between these two layers of gradedness can be difficult to grasp: in section 4.2 and 4.3 I will explain in detail this subtle distinction. In what follows, I introduce two counterexamples to the dichotomic view: namely, lies that involve beliefs about graded truth values (4.2) and lies that involve graded beliefs (4.3). I develop a non-dichotomic alternative to the TIC that counts this kind of statement as lies and allows for various degrees of insincerity in lying, according to which the speaker must believe his statement to be *more likely to be false than true* (4.4).

⁶⁶ A third option sometimes considered is that the speaker has *no opinion about p* (he *lacks a credal state about p*); I will come back on this in section 4.3.

⁶⁷ In epistemology, “credal state” indicates a specific kind of mental state: *i.e.* the mental state of having a belief. Similarly, “credence” denotes a belief, in particular a graded belief (cf. section 4.3).

4.2 Graded Truth Values

Every species is vague, every term goes cloudy at its edges, and so in my way of thinking, relentless logic is only another name for stupidity—for a sort of intellectual pigheadedness.

H.G. Wells, First and Last Things (1908)

A first motive to challenge the dichotomic view emerges if one considers the question of lying from outside the framework of bivalent (*i.e.* two-valued) logics. According to traditional, bivalent logics, the truth value of a proposition is either true or false, *tertium non datur*, so that there is no point in distinguishing between a statement that is false and a statement that is not true. By contrast, many-valued logics allow for a larger set of truth values. If ‘true’ and ‘false’ are not the only two possible truth-values that one can assign to propositions, the assumption that speakers can only have beliefs involving these two truth-values strikes as spurious, or at least unwarranted.

In the literature on lying, Chisholm & Feehan have offered a discussion of insincerity that can be read as a challenge to a bivalent account of insincerity⁶⁸. According to their alternative insincerity condition, a speaker lies only if he states what he believes to be *false* or *not true* (I call this the *fuzzy insincerity condition*, FIC).

Fuzzy Insincerity Condition

FIC: S believes p to be false or not true

This alternative formulation of the insincerity condition relies on the subtle difference between believing that *p* is *false* and believing that *p* is *not true*. Chisholm & Feehan (1977, 152) note that “it is logically possible to believe one of these things [e.g. *p* is not true] and not the other [e.g. *p* is false]”. One way to interpret this claim is to consider the difference

⁶⁸ Chisholm & Feehan do not discuss their alternative insincerity condition in detail, nor they mention explicitly that their aim is to challenge a bivalent account of beliefs: the “challenge” I mention here is thus quite indirect. My primary aim is not an exegesis of their article; I merely take a cue from their work to develop an alternative insincerity condition that allows for degrees of insincerity.

between false and not true within a specific many-valued logic theoretical framework, namely *fuzzy logic*⁶⁹.

Fuzzy logic is a many-valued logic conceived especially for predicates that are intrinsically vague (like being bald, or old, or happy) and that, being graded in nature, allow for a number of truth values. Fuzzy logic takes as truth values all real numbers between 0 and 1, where 0 is false and 1 is true. From this perspective, to say that the speaker believes that a proposition p is *not true* is to say that the speaker believes that the truth value of p is x , where x is $0 \leq x < 1$. By contrast, to say that the speaker believes that a proposition p is *false* is to say that the speaker believes that the truth value of p is 0. Against the dichotomic view, stating what is *believed to be not true* is thus not the same as stating what is *believed to be false*. Believed-false statements are a subset of insincere statements.

Let us call lies that involve these intermediate beliefs (beliefs about graded truth values) *fuzzy lies*. Now, consider an example of *fuzzy lie* to grasp the difference from traditional lies. Suppose that Mickey utters (1) to persuade Daisy to date Donald:

(1) Donald is not bald

If Mickey thinks that Donald is *almost definitely* bald (e.g. he believes that (1) has a truth value of 0.2), he does not say something that he believes to be utterly false, and therefore the TIC does not count (1) as lying. However, intuitively Mickey is lying⁷⁰. The fuzzy insincerity condition FIC accommodates our intuitions in counting (1) as a *fuzzy lie*, since (1) is believed to be *not true* (0.2 is less than 1 and more than 0).

FIC is broader than TIC: it allows all statements whose truth value is believed to be less than 1 to count as lies. Moreover, it correctly rules out misleading statements (given that they are believed to be true), while ruling in all standard lies (given that they are believed to be false). Nevertheless, it seems patent that the FIC is too broad, for the set of statements

⁶⁹ Other interpretations of the claim are possible, but they will not be discussed here, since the aim of this chapter is to outline the graded dimensions of lying. For a broader discussion of many-valued and fuzzy logics, see Hajek (1998) and Gottwald (2001, 423-492).

⁷⁰ Since I am focusing on the insincerity condition, I will always assume that in my examples the other conditions for lying obtain (i.e., that p is asserted with the intention to deceive). One might object that in this example (and in some of the following) condition (iii) does not obtain, because the speaker does not believe that his statement is utterly false, and thus does not believe that the statement is utterly deceiving. However, several philosophers (e.g. Chisholm & Feehan 1977, 145; Fallis 2011, 45; Staffel 2011, 301) argue that intending to alter someone's degree of belief counts as intending to deceive them. Moreover, I have already provided strong reasons to doubt that the intention to deceive is a necessary condition for lying.

it allows to count as lies is too large. For instance, if Mickey believes (1) to have a truth-value of 0.8, FIC would predict that Mickey is lying, but this is counterintuitive, as Mickey in this case believes that Donald is *almost definitely* not bald. To avoid this problem, one could narrow the FIC in order to require that the believed truth value of the statement be closer to falseness than truthfulness- make it x , where x is $0 \leq x \leq 0,5$ (call this the *revised fuzzy insincerity condition*, henceforth FIC*).

Fuzzy insincerity condition, revised

FIC': S believes that p has a truth value comprised between 0 and 0,5*

This solution is nevertheless problematic, since there seems not to be a clear theoretical basis to set the limit at a precise value. If one accepts that a speaker who believes that the truth value of his statement is 0.5 is lying, then it also seems reasonable to accept that a speaker who believes that the truth value of his statement is 0.51 is lying. But the same line of reasoning would work for the successive values (0.52; 0.53; [...]; 1), so that, in the end, all statements would count as lies.

A further problem with this view concerns the very existence of such credal states: the proposed representation of the speaker's beliefs is so fine-grained that it may appear to overrate reality. It seems quite clear that in real life we do not experience the threshold between believing a statement to have a truth value of 0.50 rather than 0.51. If such a threshold exists, it is not consciously perceived; and since lying is a conscious choice, no such threshold can be taken as a necessary condition for lying.

Rather than positing a sharp 'numerical' threshold that separates insincerity from sincerity, one could limit the FIC to require that the believed truth value of the statement is *perceivably* closer to falseness than truthfulness.

Fuzzy insincerity condition, revised again

FIC'': S perceives p's truth value to be closer to falseness than truthfulness

The revised condition FIC'' may seem rough compared to its 'numerical' translation, but it acknowledges that our beliefs only roughly (and rarely) correspond to the subtle differences that fuzzy logic outlines. Whether or not one finds this revised definition convincing, eventually we will be forced to abandon it: as I will show in the next section, it incorrectly rules out lies that involve graded beliefs.

4.3 Graded beliefs

The dichotomic view holds that we either believe something to be true, or we believe something to be false. This is certainly true if we restrict our analysis to cases of certainty. Here, by certainty I am referring to what philosophers call ‘psychological’ certainty⁷¹: the highest degree of confidence that a subject can have in the truth of a proposition. As long as this state of mind is concerned, it is certainly true that one can only be supremely confident that a proposition is true, or supremely confident that it is false.

It does not seem, however, that we can only be supremely confident in the truth or falsity of a proposition. Quite the contrary: many of the beliefs we hold in our daily life involve a certain degree of uncertainty. This prompts the question of which kind of belief is involved in cases where certainty is not present. If certainty is the highest degree of confidence, there must be beliefs that involve a degree of confidence lower than certainty. These weaker beliefs, that do not fit within the framework of the dichotomic view, are known in the literature as ‘credences’, or ‘graded beliefs’.

That ordinary beliefs can be graded is evident if one thinks about daily situations in which a subject lacks certainty in a proposition and nonetheless, to some extent, believes that proposition. Consider some further examples: suppose Groucho holds the following three beliefs:

- (1) I have a pair of moustaches
- (2) Bulgaria will beat Azerbaijan in their next football match
- (3) There is life on some other planet in the universe.

At T1, Groucho regards (1) as certain, (2) as probable, (3) merely as more likely to be true than not. Groucho neither fully believes nor fully disbelieves (2) or (3). His partial beliefs

⁷¹ Thus understood, certainty is always relative to someone’s standpoint: it does not matter if the subject has no ground (or bad grounds) for holding that belief, because certainty only requires that the subject be supremely convinced of its truth. Philosophers often distinguish *psychological* certainty from *epistemic* certainty (Klein 1998, Reid 2008, Stanley 2008). Epistemic certainty refers to the *degree of epistemic warrant* that a proposition has, independently of the speaker’s confidence in it (*i.e.* independently of psychological certainty). While psychological certainty is purely ‘subjective’ (it only depends on the subject’s confidence), epistemic certainty is in a sense ‘objective’ (it depends on the *actual solidity* of the subject’s reasons to believe in that proposition). The literature on lying is concerned with *psychological certainty*, since the strength of the speaker’s grounds for disbelieving an assertion is irrelevant to assess whether he is insincere or not. Consequently, in this chapter, “certainty” (and “uncertainty”) will refer to *psychological* certainty (and uncertainty).

in (2) is (3) (believing to be *probable*, believing to be *unlikely*, etc.) are what philosophers call ‘graded beliefs’, because they can be ordered in a graded scale⁷²: Groucho is more confident in the truth of (1) than he is in (2), and in (2) than he is in (3). Formal accounts of degrees of belief (namely Bayesian accounts) represent this scale with real numbers from 0 to 1, where 0 indicates certainty in the falsity of p , 1 indicates certainty in the truth of p , and 0.5 indicates uncertainty – in other words, that the subject regards p just as likely to be true as false. On this view, uncertainty is the middle point (0.5) of a continuum of degrees of belief whose poles are certainty in the falsity (0) and in the truth (1) of the proposition (cf. Figure 1). To provide a formal account of the previous example, one could say that Groucho has a degree of belief of 1 in (2), of 0.75 in (3), of 0.55 in (4).

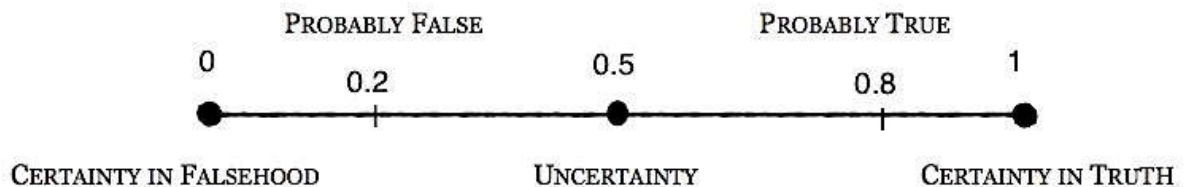


Figure 1: A visual representation of the certainty-uncertainty continuum

The fact that epistemic agents can hold a wide array of graded beliefs is at odds with the dichotomic view, that only allows for full belief in truth and full belief in falsity. Since graded beliefs and uncertainty are ordinary psychological states, it seems that a theory of lying should account for them (Meibauer 2014: 223, D’Agostini 2012:41). For instance, suppose that Groucho states (2) (that Bulgaria will beat Azerbaijan) while believing that it is probably false, or as likely to be false as true. Would his statement be sincere or insincere? More generally, how are we to draw the boundary between sincere and insincere utterances, and (consequently) between lies and not lies?

4.4 A ‘graded’ definition of insincerity

To see that the standard account of insincerity struggles to handle graded beliefs in a satisfactory way, let us consider a new example, inspired by recent historical events (cf.

⁷² For a discussion of the mutual relations between flat-out beliefs and graded beliefs, see Frankish (2009).

Carson 2010:212-21): George is a political leader, and tells (1) to a journalist. Propositions (a), (b), and (c) indicate George's degree of confidence in his utterance, in four possible scenarios⁷³.

(1) Iraq has weapons of mass destruction

- (a) $(1/\neg p)$ [Iraq has *certainly no* weapons of mass destruction]
- (b) $(0.75/p)$ [*Probably*, Iraq has weapons of mass destruction]
- (c) $(0.75/\neg p)$ [*Probably*, Iraq *does not* have weapons of mass destruction]

Scenario (1a) is a clear-cut case of lying, since George believes (1) to be *certainly false*: the traditional insincerity condition (TIC) correctly tracks the intuition that, since George believes (1) to be false, (1) is a lie. In (1b), by contrast, George believes the statement to be *probably true*: even if he is not completely confident that the statement is true, it seems that in this case he is not lying (but cf. Marušić 2012:8). The utterance is inaccurate, and perhaps misleading, because it misrepresents George's degree of belief in (1). However, being inaccurate or misleading is clearly not the same as lying (Saul 2012, Stokke 2013b). Also in this case, TIC makes the right predictions.

Problems arise for scenario (1c), where George believes (1) to be *probably false*. It seems that TIC does not count this case as a lie, because George does not utterly believe (1) to be false⁷⁴. However, intuitively this is a case of lying, because George is saying something he believes to be very likely false. Since it excludes this sort of cases, TIC is too narrow, and needs some refinement.

Cases like (1b,c) suggests that a more fine-grained account of lying is needed, one that appreciates how lying can involve graded beliefs. The fuzzy insincerity condition (FIC') will be of little help here. Such condition accounts for lies that involve *beliefs about graded truth values (fuzzy lies)*, but it does not account for lies that involve *graded beliefs (graded-belief lies)* about plain truth values. Such a subtle difference is worth an explanation.

⁷³ Assigning a defined, numeric degree of belief to these linguistic expressions (e.g. "probably", "perhaps") merely aims to indicate how these expressions can be ordered on a scale that goes from certainty to doubt (Holmes 1982, Levinson 1983:134, Hoyo 1997). Only their reciprocal relation in the scale matters to the present discussion – the accuracy of the numeric values is not important.

⁷⁴ To save TIC against this objection, a partisan of the standard view might suggest to interpret TIC in a non-literal sense, so that (2) counts as a case of believing p to be false, and hence as lying. However, this broad interpretation would open the problem of which intermediate credal states count as believing false and which do not. Since this is exactly the problem that the sincerity condition should solve, TIC would still be an unattractive option to settle the issue.

Let us represent the general structure of beliefs as “ $B(p)$ ”, where the variable “ B ” takes beliefs as values, and the variable “ p ” takes the truth-value of the propositional content of beliefs as values. The dichotomic view assumes that both “ p ” and “ B ” can assume as values only 1 or 0: either a subject believes p , or he does not believe p , and either he believes p to be true, or he believes p to be false.

Non-dichotomic accounts, by contrast, assume that “ p ” and/or “ B ” can take as values all the real numbers from 0 to 1. *Fuzzy lies* involve non-whole “ p ” values, while *graded-belief lies* involve non-whole “ B ” values. In 4.2 I provided an example of a *fuzzy lie*; let us contrast it with the *graded-belief lie* now:

- (1) Iraq has weapons of mass destruction
- (2) Donald is not bald

If George is confident, but not sure that (1) is false (*e.g.*, a degree of confidence of 0.2 in (1)), it seems clear that George is lying: his lie is a *graded-belief lie*. In this case, 0.2 expresses the value of “ B ”, so that 0.2 indicates George’s subjective degree of confidence in the probability of (1). This case is different from the *fuzzy lie* example discussed in 4.3: in that case, 0.2 indicated the truth value of (2). In the fuzzy lie example, Mickey had an outright belief ($B=1$) that Donald is almost definitely bald – *i.e.*, that the truth value of (2) is 0.2.

The FIC allows for fuzzy lies like (2) to count as lies, but do not count graded-belief lies like (1) as lies, and is therefore too narrow. An alternative insincerity condition that allows *graded-belief lies* can be found in Carson (2006: 298). His proposal comes in two varieties: he presents a strong and a weak version of his ‘insincerity condition’ for lying. The first, ‘strong’ version requires that the speaker believe his assertion to be “false or probably false”. Let us call Carson’s first condition the ‘strong insincerity condition’ for lying (SIC):

Strong insincerity condition

(SIC) S believes p to be at least probably false⁷⁵

SIC correctly captures prototypical cases of lying like (1a) (repeated below). Unlike the traditional definition, it also includes lies that are not believed with certainty to be false, like (1c), that George believes to be *probably false*. This is an advantage of SIC over the

⁷⁵ I rephrased Carson’s condition to avoid the counterintuitive consequence that degrees of belief included between “believing false” and “believing probably false” would not count as lies.

traditional condition TIC, since it seems intuitive that saying what you believe to be *probably* false counts as lying – even if it is arguably less insincere, and less deceptive, than a full-fledged lie.

However, the limit set by the SIC strikes as arbitrary: it is not clear what justifies drawing the boundary between sincerity and insincerity exactly on the degree of confidence indicated by ‘probably’, and not someplace else. The term ‘probably’ indicates a degree of confidence in the proposition higher than uncertainty and lower than certainty: for the sake of the argument, let us assume it stands for a degree of belief of 0.75 or higher. If a degree of belief of 0.75 in the falsity of the proposition is enough for lying, there seems to be no reason to exclude lower graded beliefs like 0.7, or 0.6, that are perceivably higher than uncertainty (0.5).

(1) Iraq has weapons of mass destruction

- (a) $(1/\neg p)$ [Iraq has *certainly no* weapons of mass destruction]
- (c) $(0.75/\neg p)$ [*Probably*, Iraq *does not* have weapons of mass destruction]
- (d) $(0.6/\neg p)$ [*Presumably* Iraq *does not* have weapons of mass destruction]

For instance, in (1d), George utters what he believes to be *more likely to be false than true*, so that it seems that he is lying. However, SIC does not capture (1d), because by hypothesis George’s degree of confidence is higher than uncertainty but falls short of believing (1) to be probably false. In failing to account for the intuition that also (1d) is a lie (even if arguably less insincere than (1c)), SIC is too restrictive. Furthermore, it is not clear that Carson’s SIC is able to capture fuzzy lies: only outright falsity is mentioned in the formulation, leaving it unspecified whether believing a proposition to have a truth value of more than 0 would qualify to satisfy SIC.

Carson’s second, ‘weak’ proposal avoids both problems. The ‘weak insincerity condition’ (WIC) posits that lying requires that the speaker “does not believe [the asserted proposition] to be true” (Carson 2006, cf. also Davidson 1985:88, Sorensen 2007:256, 2011:407, Fallis 2013:346).

Weak insincerity condition

(WIC) S does not believe p to be true

Since it acknowledges that utterances like (1d) are lies, WIC is preferable to SIC. Furthermore, WIC seems compatible with fuzzy lies. Arguably, if S believes *p* to have a

truth-value inferior to 0.5 (or perceives that value to be closer to falseness than truthfulness), S does not believe that p . If we accept this principle, WIC is broadly equivalent to FIC'. However, the WIC is too broad: it incorrectly captures cases in which the speaker has no idea whether what he says is true or false, but goes on saying it for some independent reasons. These cases are classified in the literature as bullshit (Frankfurt 1986). The typical example of bullshitter is the politician who “never yet considered whether any proposition were true or false, but whether it were convenient for the present minute or company to affirm or deny it” (Swift 1710). For instance, consider the following example of deceptive bullshitting. Nick is a politician who does not know what the acronym LGBT refers to. When asked by a journalist about his opinion on LGBT rights, Nick answers:

(2) LGBT rights are of central importance for our party

In uttering (3), Nick does not have the slightest idea whether what he said is true or false. His only concern is to trick the journalist into thinking that he knows what he is talking about. It seems that he is not lying, but the WIC counts incorrectly his statement as a lie, since Nick does not believe that his statement is true. As a matter of fact, philosophers seem to agree that, as long as the speaker has no opinion about the veracity of what he is saying, his utterance is better classified as a misleading utterance than as a lie (Saul 2012:20, Meibauer 2014: 103, but cf. Falkenberg 1988:93, Carson 2010:61-2); if one wants to account for this intuition, the WIC is too broad.

Since the SIC is too narrow and the WIC is too broad, an ideal condition has to lie somewhere in the middle. To find a middle ground between these two proposals, we can require that the speaker believe p *more likely to be false than true*. Call this the *comparative insincerity condition*:

Comparative insincerity condition

CIC: S believes p more likely to be false than true

Unlike WIC, CIC correctly rules out bullshit and statements uttered in cases of uncertainty. Unlike TIC, it counts graded-belief lies as lies. And unlike SIC, it rules in the other cases in which the speaker does not believe the statement to be true – like (1c) and (1d).

A first worry about the CIC is that it implicitly accepts the view that every belief can be represented as an assessment of probability. In case one finds this hypothesis disputable, one might prefer a phrasing that avoids a terminology that is committed to this view.

Furthermore, it might be argued that it is not clear that CIC rules in fuzzy lies. We might stipulate, in a similar spirit of what we did for WIC, that if S believes p to have a truth-value inferior to 0.5 (or perceives that value to be closer to falseness than truthfulness), S thereby believes p more likely to be false than true. But an even better solution might be to introduce a phrasing that avoids both these worries:

Comparative insincerity condition, revised

*CIC**: *S is more confident in $\neg p$ than he is confident in p*

CIC* has the same strength points of CIC, but on top of that it clearly rules in *fuzzy lies*, and is not committed to understanding graded beliefs in terms of assessment of probabilities. A last worry might survive about the very assumption that there is a clear-cut boundary between insincerity and sincerity. Perhaps there are indeterminate cases that amount to neither lying nor not lying, and we should treat insincerity and lying as vague predicates – a similar problem, after all, seemed to apply to FIC. But this intuition can be accommodated without altogether rejecting CIC: unlike other insincerity conditions, this is the only one that allows for a progressive transition from sincerity to insincerity. On the other hand, if lying and insincerity are not vague predicates and a neat point of transition is to be individuated, the CIC is fine-grained enough to identify the boundary that gets closer to our intuitions, avoiding the counterexamples to which the alternative accounts fall victim (and if one has worries similar to those that applied to FIC, CIC* can be qualified by requiring that the speaker is *perceivably* more confident in $\neg p$ than he is confident in p).

4.5 Expressing graded beliefs and graded truth values

We have seen that an assertion is insincere if there is a certain discrepancy (defined by CIC*) between *the speaker's belief* (henceforth B Ψ) and *the belief expressed by the sentence* (henceforth BA). This discrepancy can come in degrees, because of the graded nature of beliefs and of the content of beliefs – *i.e.* the graded nature of B Ψ discussed so far. For a complete picture, we need to look at the other side of the coin: how insincerity is affected by the different degrees of belief that an assertion can express – the graded nature of BA.

Speakers employ several linguistic devices to express graded beliefs, or belief about graded truth values. We know from our experience as ordinary language speakers that it is possible

to modulate the intensity of a statement, either mitigating or reinforcing it, thereby altering the strength of the belief expressed in BA . For instance, instead of simply uttering (1), a speaker can alternatively downgrade his assertion by uttering (1*) or emphasise it by uttering (1**):

(1) Giusi is pretty

(1*) Giusi is *kind of* pretty *somehow*

(1**) *Believe me*, Giusi is *absolutely* pretty

Now, the previous section has considered graded-belief lies, in which $\text{B}\Psi$ is graded. Graded assertions like (1*) and (1**), by contrast, are cases in which BA is graded: the former is intuitively a weaker assertion, whereas the latter is intuitively a stronger one. In pragmatics, the two opposite phenomena of mitigation (1*) and reinforcement (1**) have often been studied separately, and labelled with different names: for the former, “attenuation”, “weakening” and “downgrading”; for the latter, “strengthening” and “emphasising” (see Fraser 1980, Coates 1987, Bazzanella et al. 1991, Caffi 2007, Egan & Weatherson 2011). The label of intensity (Holmes 1984, Labov 1984) unifies these two opposite directions of modulation.

The intensity of an utterance can be modified along different dimensions. In what follows, I discuss how intensity markers can modify the *propositional content* of an assertion to express different graded truth-values (4.5.1); and modify the *illocutionary force* of an assertion, to express different graded beliefs (4.5.2-3).

4.5.1 Intensity and propositional content

The *propositional content* of a statement can be modulated both on the axis of quality (precision) and of quantity (augmentation or diminution). Expressions like “a little”, “very”, “much”, or “quite” are used to modify intensity on the axis of quantity. These linguistic devices allow the speaker to slightly alter the truth conditions of his statements (Lakoff 1973, 478-488). For instance, if Bruce utters (2*) rather than (2), he quantifies Robin’s gladness to a lower degree, thus altering the truth-conditions of his statement:

(2) Robin is glad

(2*) Robin is *pretty* glad

In section 4.2, I considered *fuzzy lies*, *i.e.* lies that involve *beliefs* about graded truth values. Utterances like (2*), similarly, are insincere *statements* that express *beliefs* about graded truth values. This analogy suggests that, with respect to *fuzzy lies*, we have to consider two graded layers of insincerity: the layer of the speaker's beliefs and the layer of the beliefs expressed by his statements. For instance, Bruce can tell a *fuzzy lie* either by 'plainly' stating (2) while believing that (2) is partly false (e.g. 0.3-true) or by stating (2*) while believing that (2) is utterly false (that is, believing that (2*) is partly false).

4.5.2. Two directions of belief misrepresentation

A similar, but more complex discourse applies to the *degrees of belief* expressed by assertions. Assertions that express graded beliefs are generally overlooked in the literature on lying. This is because, in standard cases, statements express a flat-out belief in the truth of the proposition, rather than a graded belief. For instance, (3) expresses a flat-out belief in the asserted proposition:

(3) Iraq has weapons of mass destruction

Not all statements, however, are as simple as (3), for some express graded beliefs. For instance, (3a) indicates that the speaker believes that (3) is probably true, and (3b) expresses uncertainty in the truth of the proposition:

(3a) (0.75/p) *Probably* Iraq has weapons of mass destruction

(3b) (0.5/p) *Maybe* Iraq has weapons of mass destruction

Few authors have raised the question of how assertions that express graded beliefs are to be analysed within a theory of lying. Meibauer (2014: 225) suggests that there are three kinds of graded insincere assertions that may qualify as lies: those "(i) expressing certainty when [you] are uncertain, those (ii) expressing uncertainty when [you] are certain, and those (iii) expressing certainty or uncertainty to a higher degree than being adequate with respect with [your] knowledge base". Since the third case seems to include the previous two, to simplify this taxonomy I will simply distinguish between two 'directions' in misrepresenting your degree of belief: namely, pretending to have a *higher degree of belief* or a *lower degree of belief* than the one you have (cf. Falkenberg 1988:93).

A first, tempting idea is to assume that these two directions are equivalent. This would mean that, from the point of view of the analysis of lying, “pretending to be more certain than you are” is *as insincere as* “pretending to be less certain than you are”. A reason to make this assumption is that the ‘discrepancy’ between your state of mind and the state of mind expressed by the statement is the same in both cases. However, at a closer look this assumption reveals it to be naïve, as the first case (*overstating*) is often perceived as being more insincere, or more misleading, than the second (*understating*). To see this, consider the two utterances:

(3c) (1/p) *Certainly* Iraq has weapons of mass destruction

(3d) (0.5/p) *Perhaps* Iraq has weapons of mass destruction

Imagine that in both cases George’s mental state is in between certainty and uncertainty, so that he believes:

(0.75/p) [Probably Iraq has weapon of mass destruction]

According to the ‘naïve’ view, (3c) and (3d) are equivalent scenarios, because the discrepancy between $B\Psi$ and $B\Lambda$ is the same (0.25). These scenarios differ only in the direction of misrepresentation: (3c) represents the speaker as having a higher degree of belief than he has, while (3d) as having a lower degree of belief. Interestingly, however, it is natural to assess (3c) as more insincere than (3d). The reason is that we tend to judge (3d) as a prudent statement, that cooperatively avoids saying more than the speaker knows, while (3c) is perceived a misleading overstatement, that the speaker has not sufficient knowledge to assert. In other words, *ceteris paribus*, understating your degree of belief is generally seen as a cooperative linguistic practice, while overstating it is generally regarded as uncooperative.

In line with this intuition, Falkenberg (1988: 94, 1990) proposes to distinguish between ‘hard lies’ (*overstatements*, like (3c)) and ‘soft lies’ (*understatements*, like (3d)). However, this taxonomy is misleading in two respects. First, not all overstatements and understatements are lies: if the CIC is a condition for lying, only statements displaying a certain level of discrepancy between $B\Psi$ and $B\Lambda$ can be lies. Second, it is not clear whether an *overstatement* (hard lie) is necessarily more of a lie than an *understatement* (soft lie): the next section will show that the direction of misrepresentation is just one of the parameters

of intensity that must be considered, another one being the magnitude of the discrepancy between $B\Psi$ and $B\Lambda$.

4.5.3. Epistemic modals and degrees of commitment

The most prominent linguistic devices used to mitigate or reinforce the degree of belief expressed by an assertion (expression like ‘certainly’, ‘probably’, ‘perhaps’) are called *epistemic modals*. This section will analyse how they alter the degree of belief expressed by the assertion, and clarify why we generally assess *understatements* as more sincere (or more honest) than *overstatements*.

On a pragmatic level, epistemic modals both “indicate the *speaker’s confidence* or lack of confidence in the truth of the proposition expressed” and “*qualify [his] commitment* to the truth of the proposition expressed in [his] utterance” (Coates 1987:112, italic is mine). In other words, they act on two components of the assertion, altering both (1) the psychological state expressed by the speaker (the degree of belief), and (2) his degree of commitment to the truth of the proposition (the illocutionary strength⁷⁶) (cf. Sbisà & Labinaz 2014:52, Lyons 1977: 793-809; Holmes 1984: 349).

These two functions are distinct in nature, but entangled: if a speaker S mitigates (or reinforces) the degree of belief conveyed by his assertion, then S automatically mitigates (or reinforces) the illocutionary force of his assertion (that is, his degree of commitment to the truth of the proposition). For instance, if you state (4b) instead of plainly stating (4), you both mitigate the degree of belief expressed ((4b) expresses uncertainty in (4)) and lower the degree of your commitment to the truth of the asserted proposition (you are committed to the truth of (4) to a much lower degree if you utter (4b))⁷⁷.

⁷⁶ The illocutionary force of an assertion can be reinforced or mitigated (Bazzanella, Caffi & Sbisà 1991; Sbisà 2000; Searle & Vanderveken 1985: 99), thus altering the speaker’s degree of commitment to the truth of the proposition. More generally, “along the same dimension of illocutionary point there may be varying degrees of strength or commitment” (Searle 1976:5). Epistemic modals and other intensity markers can modify these degrees of strength (Holmes 1984; Bazzanella, Caffi & Sbisà 1991; Sbisà 2000; Searle & Vanderveken 1985:99). For a discussion on the distinction between illocutionary and propositional mitigation, see Caffi (1999, 2007) and Fraser (2010:16-17).

⁷⁷ On this ‘*expressivist*’ interpretation, epistemic modals are not part of the proposition asserted (at least not of the proposition against which speaker sincerity and commitment is assessed). A ‘*descriptivist*’ might object that we should instead take them to be part of the content of the assertion (and hence of the proposition against which sincerity is measured). However, this would often yield counterintuitive predictions for the sincerity conditions of assertions. For instance, on a descriptive interpretation of “certainly p” as true iff (q): “the speaker is certain that p”, a speaker that believes that there are 9/10 chances that p is true would counterintuitively be counted as insincere (as S

(4) Plato will quit smoking tomorrow

(4b) *Perhaps* Plato will quit smoking tomorrow

The role that epistemic modals play in reinforcing/weakening the illocutionary force of assertions explains why *understatements* are perceived as more honest than *overstatements*. *Ceteris paribus* (given the same degree of insincerity, like in (3c)-(3d)) a reinforced assertion has a stronger illocutionary force than a mitigated assertion, so that the speaker has a stronger commitment to its truth. And if the commitment to sincerity is stronger in reinforced statements, then violating that commitment is more serious in those statements than in mitigated ones.

Variations in illocutionary force induced by epistemic modals can affect whether the speaker is asserting the proposition or not – and hence whether he is lying, because lying requires asserting. This is because epistemic modals can downgrade the degree of illocutionary force of a declarative sentence to such an extent that it no longer counts as an assertion, but rather as a supposition or a hypothesis (Sbisà & Labinaz 2014:52-3). For instance, (4b) is a supposition rather than an assertion: its insincere utterance does not amount to lying, while insincerely uttering its unmitigated version (4) does. Carson (2010: 33,38) shares this intuition: “there are weaker and stronger ways of warranting the truth of a statement. To count as a lie, a statement must be warranted to a certain minimum degree”. This is even more evident in other speech acts. For instance, if Matteo utters (5b) instead of (5), it is clear that he has not promised that he will buy you an elephant (he is merely suggesting it), while he would be promising it if he uttered (5). It seems that an insincere utterance of the first amounts to lying, while this is not true for the second⁷⁸.

would be certain that q is false). It should be noted that even if this section provides sincerity conditions for marked assertions interpreted in an expressivist fashion, it is not committed to expressivism: a descriptivist can still adopt the model proposed in section 1 (CIC). I follow Coates' (1987:130) view that epistemic modals can be appropriately used and interpreted in both ways. When they are used 'literally' to assert the epistemic or psychological (un)certainly of a proposition (rather than *express* that the proposition asserted is (un)certain, the simple sincerity conditions provided by CIC will apply; in the other cases (that I take to be the prevalent uses), the expressivist explanation outlined in this section will apply instead. On the debate over the semantics of epistemic modals, cf. Kratzer (1981), DeRose (1991), Egan, Hawthorne & Weatherston (2005), Papafragou (2006), Fintel & Gillies (2008), Yalcin (2007, 2011), Swanson (2011).

⁷⁸ One might wonder whether uttering (4b) or (3b) while being *certain* that the mitigated proposition is false would count as lying – *i.e.* if a high degree of insincerity can compensate for a low degree of commitment. Marsili (2014: 166-8) argues against this view, claiming that these utterances are to be classified as misleading statements rather than lies.

(5) Tomorrow I will buy you an elephant

(5b) *Perhaps* tomorrow I will buy you an elephant

This theoretical framework allows us to correctly analyse lying involving assertion that express graded beliefs (BA-graded lies, and complex cases where both BA and BΨ are graded). The problem with these assertions is that they cannot be dealt with simply by appealing to condition CIC. Without an account of how epistemic modals modify illocutionary force, CIC alone cannot account for the differences determined by the direction of misrepresentation (overstatements vs understatements). This difficulty dissipates once it is understood that epistemic modals influence not only whether the sincerity condition is satisfied (by altering the *degree of belief* expressed), but also whether the assertion condition is satisfied (by altering the speaker's *degree of commitment*).

To sum up, there are three entwined scalar parameters that we have to consider when we analyse the graded components of the belief expressed by an utterance: the *graded truth value* expressed by the utterance, the *graded belief* it communicates (these two parameters determine the degree of insincerity), and the *degree of assertoric force* of the utterance, or degree of commitment (this parameter determine if the speaker is asserting, and how strongly) that the speaker undertakes. For an utterance to count as a lie, both a certain degree of insincerity *and* a certain degree of commitment must obtain.

4.6 Further graded components

So far, I have considered aspects of intensity that are properly linguistic. Several paralinguistic devices can intervene to modulate the illocutionary force of an assertion. A significant factor is *prosody*, and specifically *intonation*: Gussenhoven (2002) remarks that variations in volume and tonal height are a powerful device to communicate the speakers' confidence in the truth of his assertion. Among other influential factors that can be used to influence the strength of an utterance are pauses, rhythm of speech, repetitions and proxemic signals.

Context is also an important factor to be considered. Bazzanella (2009, 78) points out that, among the parameters affecting the degree of insincerity of a lie, also “aspects of the global context [...], of the cotext [...] and of the local context” are of central importance. For instance, an utterance like (1) can express a different degree of certainty whether (1) is uttered in (context A) a restaurant, where a person addresses it to a dining companion that

is known not to like cheese, or (context B) in a hospital, where a doctor addresses it to a person that he knows to have a deadly allergy to cheese. In both cases, given that (1) is insincere, (1) is a lie, but it is more intense in the second case, since the commitment to the truth of the assertion is stronger in context A than in context B.

(1) This meal contains no cheese

The very mitigating and reinforcing devices considered so far are influenced by contexts - in most cases, their very meaning is determined by contextual factors (Kratzer 1981). For instance, the epistemic modal 'definitely' in (2) expresses a different degree of certainty and commitment depending on whether (2) is uttered in context A or in context B.

(2) Definitely, this meal contains no cheese

Linguistic, paralinguistic and contextual elements determine a complex interplay of factors that influences the degree of illocutionary force of an assertion, and thus the degree of intensity of a lie. However, it seems to me that only linguistic devices (like epistemic modals) are strong enough to affect the status of an utterance as an assertion (and therefore as a lie). Paralinguistic and contextual elements can reinforce or mitigate an assertion, but they alone cannot determine if an utterance counts as an assertion or not.

In table 1, I summarise my taxonomy of the layers of gradedness involved in lying. This taxonomy is not meant to be exhaustive, but it succeeds in underlining a fact that has been ignored by the philosophical literature on lying for a long time: the multi-layered gradedness of lying.

| | |
|---|---|
| Graded of <i>beliefs</i> | Beliefs about graded truth values |
| | Graded beliefs about truth values |
| Gradedness of <i>statements</i> | Stating “beliefs about graded truth values” |
| | Stating “graded beliefs about truth values” |
| | Modifying illocutionary strength |
| Gradedness of <i>paralinguistic components</i> | Modifying illocutionary strength |
| Gradedness of <i>contexts</i> | Modifying illocutionary strength |

Table 1: A taxonomy of the dimensions of gradedness involved in lying.

In conclusion, the gradedness of lying results from the interaction several parameters: on the side of beliefs (BΨ), graded beliefs and beliefs about graded truth values; on the side of statements expressing such beliefs (BΛ), the numerous ways to convey information about them in statements. In the literature on lying, scholars tend to completely ignore these graded features when they assume the dichotomic view as a starting point – a view that I proved to be inconsistent with our intuitions.

I have provided several reasons to believe that the traditional insincerity condition is wrong, because it rules out *fuzzy lies* and *graded-belief lies*. Moreover, I have shown that the traditional insincerity condition yields a wrong description of lying, as it blinds us to its graded nature. My proposed definition corrects this picture and allows for *fuzzy lies* and *graded-belief lies* to be counted as genuine lies, and acknowledges the existence of many degrees of insincerity in lying. I have also shown that the very modifiers that affect the degree of belief expressed by an assertion modify its *illocutionary* force, affecting whether a given utterance counts as an assertion.

My main aim in this section has been to show how the insincerity condition for lying need to be modified in order to correctly deal with cases involving graded insincerity. It should be kept in mind, however, that also other components of lying have graded features. The picture is complex: as Bazzanella (2009, 78) points out, “the different degrees of intensity in lying result from the complex interplay of various layers and parameters”. For instance, some authors (Chisholm & Feehan 1977; Fallis 2011; Staffel 2011; Marsili 2017) contend

that the intention to deceive, as well as the effects of deception, can be graded. Several pragmatic parameters determinant for lying are also graded, such as relevance (Sperber & Wilson 2002; Van der Henst et al. 2002), felicity, and the relations between the interactants (like social relations, and respective trust) (Bazzanella 2009).

Finally, moral evaluations of lying can be graded – interestingly, in a way that can parallel the graded components of lying just identified. Intuitively, some lies are worse than others: lying to save a life is better than lying to get away with murder. Many factors will intervene in determining the reprehensibility of a lie: two obvious candidates will be the *effects* of the lie and the *intentions* of liar (quite obviously, good effects and intentions are preferable to bad ones). But especially from a *deontologist* perspective (Augustine *DM*, Aquinas *ST*, Kant *GMM*, *SRTL*, Newman 1880, Geach 1977, Pruss 1999, Tollefsen 2014, cf. also Isenberg 1964), a great deal of what is wrong with lying is that lying violates a moral norm of sincerity. In pointing out that such norm can be violated to a greater or lower extent, and in developing a formal model for measuring and describing the extent of such violations, my proposed account represents also a valuable tool for assessing the (im)morality of lying.

5. CONCLUSIONS – A GENERAL ACCOUNT OF INSINCERITY

So far, in my discussion of the graded nature of insincerity I have ignored attitudes other than beliefs. In this final section, I will show how my discussion generalises to these other attitudes. As a starting point, let us reconsider the general speech-act theoretic account of insincerity presented in 2.2:

INS: The performance of an illocutionary act $F(p)$ that expresses the psychological state $\Psi(p)$ is insincere IFF in uttering $F(p)$, S is not in $\Psi(p)$

Applied to beliefs, INS gives us the following belief insincerity condition (BIC):

BIC: S asserts that p insincerely only if S does not believe that p

I have already discussed condition BIC (presented as Carson's weak insincerity condition, or WIC) in detail, showing that BIC is able to deal with cases of graded insincerity, like *fuzzy lies* and *graded lies*. However, I rejected BIC as a necessary condition for lying, on the ground that it incorrectly counts *bullshitting* (saying something you neither believe to be true nor false) as lying. Since INS implies that the insincerity condition for assertion is BIC, INS is not well suited as a general account for the insincerity conditions for lying.

On the other hand, in the IV.1 I mentioned that I am concerned with two related notions of insincerity: an account of the *insincerity condition in the definition of lying*, and an analysis of *insincerity simpliciter*. We have already seen that these two notions can come apart: for instance, speech acts that do not entail assertions (e.g. *requests*, *orders*) can be insincere but cannot be lies. INS is not viable as an account of the *insincerity conditions for lying*, but it seems an appropriate analysis of *insincerity simpliciter*. This is because bullshitting falls short of lying, but is arguably a form of insincere speech. If one shares this intuition, then bullshit is not a counterexample to BIC or INS understood as characterising *insincerity simpliciter*, rather than as a necessary condition for lying. If one, by contrast, intuitively feels that bullshitting amounts to neither lying nor insincerely speaking, the general account of the insincerity condition for lying that I am about to delineate will coincide with their desired account of *insincerity simpliciter*, as both exclude bullshitting from the definition.

How can we narrow **INS** so that it excludes bullshitting but still generalises to speech acts other than assertion? In the previous section I have argued that the comparative insincerity condition **CIC** is preferable to **BIC** as a necessary condition for lying:

*CIC**: *S is more confident in $\neg p$ than he is confident in p*

The desired refinement of **INS** must entail **CIC*** rather than **BIC**, and generalise to speech acts that express attitudes other than beliefs. To be sure, at least when it comes to insincerity conditions for lying, we do not need to show that **INS** generalises to every possible attitude. In chapter III, I have shown that only commissive and assertive illocutionary acts can entail an assertion and thus count as lies (III.4). Since these families of speech acts only express either *beliefs* or *intentions* (Searle 1976), a general account of the insincerity conditions for lying only needs to generalise to intentions and beliefs. In other words, our minimal desideratum is an extension of the graded insincerity conditions for beliefs (**CIC***) to intentions.

It seems that we can reformulate the intention insincerity condition (**IIC**) in a way that parallels **CIC***, namely in a way that involves a comparison between a psychological state and its opposite:

IIC: *S does not intend to p*

*IIC**: *S intends to not- p more than S intends to p*

Admittedly, the **IIC*** is a bit odd-sounding. One reason is that intentions do not seem to come in degrees in the same way belief do (but cf. Holton 2008). Intuitively, intending is an ‘on/off predicate’: either one intends to eat an apple, or one does not intend to eat an apple – it is just not clear which intermediate mental state could exist between the two. If this is right, it is not clear how allowing for intermediate cases between an intention and its opposite can introduce a meaningful refinement of **IIC**.

Perhaps **IIC*** can be interpreted as involving a comparison between graded truth-values, rather than graded intentions. In other words, the point of **IIC*** would be to capture promises that express insincere intentions akin to *fuzzy lies*, rather than *graded-belief* lies. Suppose for instance that I promise to my girlfriend:

(1) I promise (1*) [that I will get fit]

In promising (1), I express an intention with content (1*), namely an intention to get fit. It seems that this intention can be insincere to different extents: I can intend to definitely get fit (in which case (1*) has value of 1), intend to definitely not get fit (in which case (1*) has a value of 0), and I can have a number of intermediate intentions, depending on the graded truth-value assigned to (1*) – for instance, intending to get *somewhat* fit (in which case (1*) would have a value of, say, 0.3), get *quite* fit, etc. Unlike IIC, IIC* allows for these intermediate states. On top of this, it individuates a plausible boundary between these intermediate states, for the same reason that CIC* individuates the right boundary for beliefs – namely, it seems that the limit between insincerity and sincerity does not lie at any arbitrary point close to the extremes, but rather in between them.

This is good news, as it shows that the CIC* can be neatly extended to attitudes other than beliefs. We can thus derive the following general insincerity condition for lying, from which both IIC* and CIC* can be derived:

Graded insincerity condition for illocutionary acts

INS-L: The performance of an illocutionary act $F(p)$ is insincere IFF in uttering $F(p)$, S is in $\Psi(\neg p)$ more than S is in $\Psi(p)$

INS-L is designed to deal with cases involving intentions or beliefs, but it also generalises to other attitudes, such as desires. If one has the intuition that *bullshitting* is not a form of insincerity, INS-L will also offer a plausible characterisation of insincerity *simpliciter*. More importantly, INS-L offers the general background that is needed to offer a general definition of lying that applies to speech acts other than assertions and attitudes other than beliefs. Such a definition, derived by integrating INS-2 into the definition of lying developed in chapter III, reads as follows:

Speech act theoretic definition of lying

In successfully uttering an illocutionary act with content p that expresses an attitude $\Psi(p)$ and entails an assertion with content p , S lies to A about p iff:

1. *S thereby asserts that p (i.e. conditions (a), (b), and (c) from III.3.1 obtain)*
2. *Either S is more confident in p than S is confident in $\neg p$, or S is in $\Psi(\neg p)$ more than S is in $\Psi(p)$, or both*

In the case in which the speaker is asserting directly, this definition will reduce to the following:

Definition of lying by asserting

In successfully uttering an illocutionary act with content p , S lies to A about p iff:

1. *S thereby asserts that p (i.e. (a), (b), and (c) obtain)*
2. *S is more confident in $\neg p$ than S is confident in p*

This completes this dissertation's reflection on the definition of lying. My proposed definition is perhaps complex and thus less elegant than competing ones, but it gives better predictions than the alternative accounts, and finds independent theoretical support from a general theory of illocutionary acts. Unlike deceptionist accounts, it is not subject to counterexamples to the intention to deceive condition. Unlike assertion-based accounts, it deals correctly with lies involving explicit performatives. And unlike any other account, it is able to capture to speech acts other than assertions and attitudes other than beliefs, and to include both *graded-belief lies* and *fuzzy lies*.

Thus far, I have provided a characterisation of what it means to lie and to be insincere. Avoiding lying and insincerity, however, is not all that we require from our interlocutors. In the next chapters, I will deal with another expectation that is fundamental for our communicative interactions, namely expectations about the epistemic standpoint of our interlocutors. When someone asserts something, we generally expect such claim to be backed up by some reasons, rather than mere belief – in other words, all things the same, a 'gut feeling' that a proposition is true is not enough for it to be assertable. However, it is not clear exactly which kind of epistemic standpoint is required for an assertion to be permissible *qua assertion*. The next chapter will deal the 'epistemic norm of assertion' hypothesis, namely the idea that assertions are subject to a norm of the form: "assert that p only if p has C ", where C indicates the unique epistemic property that a proposition must have for it to be assertable.

Part B - The norms of assertion

B. The norms of assertion

Quisquis autem hoc enuntiat quod vel creditum animo, vel opinatum tenet, etiamsi falsum sit, non mentitur. Hoc enim debet enuntiationis suae fidei, ut illud per eam proferat, quod animo tenet, et sic habet ut profert. Nec ideo tamen sine vitio est, quamvis non mentiatur, si aut non credenda credit, aut quod ignorat nosse se putat, etiamsi verum sit: incognitum enim habet pro cognito.

Now whoever utters that which he holds in his mind either as belief or as opinion, even though it be false, he lies not. For this he owes to the faith of his utterance, that he thereby produce that which he holds in his mind, and has in that way in which he produces it. Not that he is without fault, although he lie not, if either he believes what he ought not to believe, or thinks he knows what he knows not, even though it should be true: for he accounts an unknown thing for a known.

Augustine, De Mendacio, III.3

In the previous chapters, I have provided a detailed analysis of a common communicative vice: lying, understood in terms of asserting something that you believe to be false. In this chapter, I will discuss another kind of communicative vice: that of asserting something in absence of the appropriate warrant or epistemic support.

Once again, Augustine's *De Mendacio* offers an interesting starting point to introduce the debate: in the passage quoted above, we can find what is arguably the first philosophical discussion of assertions that are sincere, but nonetheless unwarranted. Here Augustine is discussing the distinction between *inadvertently saying something false* and *lying*. We have already seen that this distinction is intuitively correct: lying and being mistaken about what you say are two different concepts. Importantly, here Augustine finds it necessary to specify that while *inadvertently saying something false* falls short of lying, also this kind of assertion involves a *vitium* - a fault, violation or vice.

What kind of *vitium* or fault exactly? There are many faults that we might identify in such assertions: lack of epistemic support, or more simply lack of accuracy, or of correspondence with reality. To make the discussion more concrete, let us consider an example. Suppose that Claudia and Rachel are having a conversation about their common friend Jacques, and Rachel asks Claudia whether Jacques is a good cook. Claudia has never tasted Jacques' cuisine, nor has she got any second-hand information about Jacques' abilities in the kitchen. However, Claudia knows that Jacques is French, and she is under the impression that French people are generally good cooks. On this basis, she replies:

(1) Sure, Jacques is a good cook

Suppose that Jacques is not, as a matter of fact, a good cook: he is utterly terrible in the kitchen. In this case Claudia is not lying; nonetheless, her assertion is not "without fault", as Augustine would say. But which kind of fault is exactly involved in Claudia's assertion that (1)?

A contemporary approach to explaining what is wrong with (1) is to point out that the speaker (in this case, Claudia) failed to meet some relevant conversational norm, and consequently some conversational expectation. Arguably, assertors should have some epistemic ground to support what they say: there seem to be an implicit norm that dictates that they should meet some minimum epistemic standard before they assert something. In our example, (1) is false, and Claudia is not justified to believe that (1) is true. One way to explain what is wrong with (1), then, is to say that in uttering it Claudia violates a putative conversational norm – for instance, a norm requiring that you only utter true (or known, or reasonably believed) propositions. But Claudia's assertion could also violate some other conversational expectation. In asking whether Jacques is a good cook, Rachel is expecting Claudia to base her assertion on some appropriate grounds, not on a wild guess; she expects her to assert a true proposition, not a false one. At least on a first intuitive level, it seems a reasonable hypothesis that we perceive wild guesses like (1) to be faulty because they violate some relevant conversational norm, or at least the expectations that might be reasonably be held by the participants in the conversation.

In the last twenty years, the hypothesis that the faultiness of assertions like (1) might be explained in terms of the violation of a putative 'norm of assertion' has gained centre stage in philosophy of language and epistemology, sparking a lively debate around the question of which norm regulate assertions. Timothy Williamson's "Assertion" (1996, revised in

2000) is the perhaps the first paper to explicitly and systematically address this question, and definitely the one responsible for initiating the contemporary debate over this issue. Williamson opens his paper by putting forward a simple hypothesis (Williamson 2000:241):

WILLIAMSON'S HYPOTHESIS

What are the rules of assertion? An attractively simple suggestion is this. There is just one [constitutive]⁷⁹ rule. Where C is a property of propositions, the rule says:

(The C-rule) • One must: assert p only if p has C.

According to this hypothesis, there is only one rule, the C-rule, to which all and only assertions are subject – a rule that tells you which propositions you can properly assert and which ones you cannot. This rule requires you to assert only propositions that have the unspecified property 'C'.

Williamson clarifies that the C-rule (1) is *unique* to assertion: *only* assertion is regulated *only* by this rule. Since only assertions are regulated only by this rule, the C-rule also (2) *individuates* assertion, defining it as the only speech act that is only subject to the rule. Furthermore, the C-rule (3) is *constitutive* of assertion: were assertion regulated by a different rule, it would be a different speech act. Furthermore, and relatedly, the C-rule is (4) a norm to which assertors are subject *qua assertors*. There are many normative constraints that can contribute in making an assertion *overall* wrong: an assertion can be inappropriate because it is impolite, immoral, rude or irrelevant. While these are all good reasons not to proffer an assertion, they are not infractions that of a norm to which *only* assertions are subjects: commands, questions (and perhaps even some non-communicative behaviours) can be impolite, immoral, rude or irrelevant too. The norm that Williamson is looking for, by contrast, is a norm to which *only* assertions are subject, and to which assertors are subject *in virtue* of their being asserting something.

WILLIAMSON'S HYPOTHESIS is indeed "attractively simple": if correct, we can both define what assertion is and specify under which condition its performance is appropriate, simply by identifying one property, C. This raises the question that animates the debate: what kind of property is C? Williamson's answer is that C is the property of *being known by the*

⁷⁹ I have incorporated the claim that the norm is constitutive into the quotation for simplicity: in this way, all of Williamson's key assumptions are displayed in a single passage.

assertor. In other words, assertion is governed by the norm that one should not assert what one does not know to be true:

WILLIAMSON'S ANSWER: THE KNOWLEDGE-RULE

KR: "You must: assert that *p* only if you know that *p*".

A number of philosophers have found *KR* a convincing answer (e.g. DeRose 2002; Hawthorne 2004; Benton 2014). However, the debate is far from settled, and Williamson's position has elicited a number of critical responses. One could roughly divide these critical reactions into two categories.

The first category comprises those who reject or challenge WILLIAMSON'S HYPOTHESIS, either in part or as a whole. Some philosophers (Brown 2008, Carter 2014; 2017, Carter & Gordon 2011, Gerken 2014, McKenna 2015) reject the assumption that there is *only one* norm of assertion. Others deny that assertion is regulated by a norm (Sosa 2009, cf. also Rescorla 2007; 2009), or that it is *constituted* by it (Hindriks 2007, Maitra 2011, Pagin 2011, McCammon 2014:137-9). Finally, some have gone so far as to claim that there is no such thing as an 'assertion-game' to which the putative rule applies (Cappelen 2011, Johnson 2017).

In the next chapter, I address criticisms of this kind. More specifically, I object to the claim that the C-rule is 'constitutive' of assertion. After reviewing several difficulties for this hypothesis, I show that abandoning the idea that the norm of assertion is constitutive also puts strain on the assumption that the norm is *unique* to assertion, and consequently on the assumption that assertion can be defined as the only speech act subject to this rule.

The second category of criticisms comprises those coming from scholars who accept WILLIAMSON'S HYPOTHESIS, but refuse WILLIAMSON'S ANSWER in favour of a different one, *i.e.* an alternative account of what property C is. For instance, some maintain that a warranted assertion requires instead the *truth* of the proposition (Weiner 2005, Whiting 2012), or some relevant *reason to believe it* (Lackey 2007, Kvanvig 2009).

In chapter VI, I consider criticisms of this kind. More specifically, I address the disagreement between factive and non-factive accounts, *i.e.* accounts that do (like Williamson's knowledge-rule) or do not require that the proposition is true. After presenting some objections to factive accounts, I present my own non-factive proposal. On this view, false assertions are faulty because they fail to meet the success-condition for asserting something (the purported aim of assertion), rather than a permissibility condition (the norm regulating assertion).

References

- Adler, Jonathan E. 1997. Lying, Deceiving, or Falsely Implicating. *Journal of Philosophy*. 94: 435-452.
- ALDRICH, VC. 1966. Telling, Acknowledging and Asserting. *Analysis* 27 (2): 53-56.
- Alicke, Mark, David Rose & Dori Bloom. 2014. Causation, norm violation, and culpable control. In J. Knobe & S. Nichols (eds.). *Experimental Philosophy* (vol. 2), 229-250. New York: Oxford University Press.
- Alston, WJ. 2000. *Illocutionary Acts and Sentence Meaning*. Ithaca: Cornell University Press
- Anscombe, Elizabeth. 1981. *Ethics, Religion and Politics*. Oxford: Blackwell.
- Aquinas. [ST] *Summa Theologiae*
- Arico, Adam J., & Fallis, Don. 2013. Lies, Damned Lies, and Statistics: An Empirical Investigation of the Concept of Lying. *Philosophical Psychology* 26 (6): 790-816.
- Augustine. [DDC] *De Doctrina Christiana*
- Augustine. 1887 [DM]. *De mendacio* [On lying]. In *Nicene and Post-Nicene Fathers, First Series, Vol. 3*. ed. by Philip Schaff. Buffalo, NY: Christian Literature Publishing Co.
- Austin, John L. 1961/2003. *Performative Utterances*. In J.L. Austin, *Philosophical Papers*. Oxford: Oxford University Press, 3rd edn.
- Austin, John L. 1962/1975. *How to Do Things with Words*, Oxford: Oxford University Press, 2nd edn.
- Bach, Kent and Robert M. Harnish. 1979. *Linguistic Communication and Speech Acts*. Cambridge, Mass.: MIT Press.
- Bach, Kent. 1975. Performatives are statements too. *Philosophical Studies* 28: 229-236.
- Bach, Kent. 2007. Knowledge in and out of context. In Campbell, J. K. & O'Rourke, M. (eds.), *Knowledge and skepticism*. Cambridge: MIT Press.
- Bach, Kent. 2008. Applying Pragmatics to Epistemology. *Philosophical Issues* 18 (1): 68-88.
- Barnes, John A. 1994. *A Pack of Lies: Towards A Sociology of Lying*. Cambridge, NY: Cambridge University Press.
- Bazzanella, Carla, Caffi, Claudia & Sbisà, Marina. 1991. Scalar dimensions of illocutionary force. In Igor Ž. Žagar (ed.), *Speech acts: fiction or reality* (63-76). Antwerp/Ljubljana: IPrA Distribution Center for Yugoslavia and Institute for Social Sciences.
- Bazzanella, Carla. 2009. Approssimazioni pragmatiche e testuali alla menzogna. (Pragmatic and textual approaches to the concept of lying). In *Tra Pragmatica e Linguistica testuale. Ricordando Maria-Elisabeth Conte*, Venier Federica (ed.), 67-90. Dell'Orso: Alessandria.
- Beardsley, Monroe C. 1981. *Aesthetics: Problems in the Philosophy of Criticism*. Indianapolis: Hackett Publishing (Second Edition, 1958).
- Belnap, Nuel D. 2000. Double Time References: Speech-act Reports as Modalities in an Indeterminist Setting. *Advances in modal logic* 3: 37-58.

- Benton, Matthew A. 2013. Dubious Objections from Iterated Conjunctions. *Philosophical Studies* 162 (July 2011): 355–58. doi:10.1007/s11098-011-9769-3.
- Benton, Matthew A. 2014. Gricean Quality. *Nous* 50 (4): 689–703. doi:10.1111/nous.12065.
- Benton, Matthew A. forthcoming. Lying, Belief, and Knowledge. In J. Meibauer, *The Oxford Handbook on Lying*. Oxford: Oxford University Press.
- Boghossian, Paul A. 2003. The normativity of content. *Philosophical Issues* 13: 31–45.
- Bok, Sissela. 1978. *Lying: Moral Choice in Public and Private Life*. New York NY: Random House.
- Boyd, Kenneth. 2015. Assertion, Practical Reasoning, and Epistemic Separabilism. *Philosophical Studies* 172 (7): 1907–27. doi:10.1007/s11098-014-0378-9.
- Brandom, Robert. 1983. Asserting. *Nous* 17 (4): 637–650.
- Brandom, Robert. 1994. *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge: Harvard University Press.
- Brown, Jessica. 2008. The Knowledge Norm for Assertion. *Philosophical Issues* 18(1): 89–103.
- Caffi, Claudia. 1999. On Mitigation, *Journal of pragmatics* 31: 881–909.
- Caffi, Claudia. 2007. *Mitigation*. London: Elsevier.
- Cappelen, Herman. 2011. Against Assertion. In Brown, J. & Cappelen, H. (Eds.). *Assertion: New Philosophical Essays*. Oxford: Oxford University Press.
- Carson, Thomas L. 1988. On the Definition of Lying: A reply to Jones and revisions. *Journal of Business Ethics* 7:509–514.
- Carson, Thomas L. 2006. The definition of lying. *Nous* 2: 284–306.
- Carson, Thomas L. 2009. Lying, deception, and related concepts. In *The philosophy of deception*, ed. Clancy Martin, 153–87. New York: Oxford University Press.
- Carson, Thomas L. 2010. *Lying and Deception*. Oxford: Oxford University Press.
- Carson, Thomas L., Richard E Wokutch and Kent F Murrmann. 1982. Bluffing in Labor Negotiations: Legal and Ethical Issues. *Journal of Business Ethics* 1(1): 13–22.
- Carter, J Adam. 2017. Assertion, Uniqueness and Epistemic Hypocrisy. *Synthese* 194 (5): 1463–76. doi:10.1007/s11229-015-0766-5.
- Carter, J. Adam. 2017. Assertion, Uniqueness and Epistemic Hypocrisy. *Synthese* 194 (5): 1463–76. doi:10.1007/s11229-015-0766-5.
- Carter, JA. & EC. Gordon. 2011. Norms of Assertion: The Quantity and Quality of Epistemic Support. *Philosophia* 39(4): 615–635.
- Chan, T., & Kahane, G. 2011. The trouble with being sincere. *Canadian Journal of Philosophy* 41(2): 215–234.
- Chan, Timothy, and Guy Kahane. 2011. The Trouble with Being Sincere. *Canadian Journal of Philosophy* 41 (2): 1–13.
- Chen, Rong, Chunmei Hu and Lin He. 2013. Lying between English and Chinese: An Intercultural Comparative Study. *Intercultural Pragmatics* 10.3: 375–401.
- Chisholm, Roderick M. & Feehan, Thomas D. 1977. The intent to deceive. *Journal of Philosophy* 74(3): 143–159.

- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Coates, Jennifer. 1987. Epistemic modality and spoken discourse. *Transactions of the Philological Society* 85(1): 110-131.
- Coffman, E. J. 2014. Lenient accounts of warranted assertability. In C. Littlejohn & J. Turri (Eds.), *Epistemic norms: new essays on action, belief and assertion* (33-59). Oxford: Oxford University press.
- Cole, Shirley A. N. 1996. Semantic Prototypes and the Pragmatics of *Lie* Across Cultures. *The LACUS Forum* 23:475-83.
- Coleman, Linda and Paul Kay. 1981. Prototype Semantics: The English Verb 'lie'. *Language* 57: 26-44.
- D'Agostini, Franca. 2012. *Menzogna* (Lying). Torino: Bollati Boringhieri.
- Davidson, D. 2001. *Inquiries Into Truth and Interpretation*. New York: Oxford University Press
- Davidson, Donald. 1985. Deception and Division. In J. Elster (ed.), *The Multiple Self*. Cambridge: Cambridge University Press.
- Dawkins, Richard. 1989. *The Selfish Gene* (2nd ed). Oxford: Oxford University Press.
- DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of personality and social psychology*, 70(5), 979.
- DeRose, K. 2002. Assertion, Knowledge, and Context. *The Philosophical Review* 111(2): 167-203.
- DeRose, Keith. 1991. Epistemic possibilities. *The Philosophical Review* 100(4): 581-605. DOI: 10.2307/2185175.
- Donagan, Alan. 1977. *A Theory of Morality*. Chicago: Chicago University Press.
- Douven, Igor. 2006. Assertion, Knowledge, and Rational Credibility. *Philosophical Review* 115 (4): 449-85. doi:10.1215/00318108-2006-010.
- Douven, Igor. 2009. Assertion, Moore, and Bays. *Philosophical Studies* 144 (3): 361-75. doi:10.1007/s11098-008-9214-4.
- Dummett, M. 1981. Assertion, in *Frege: Philosophy of language*. London: Duckworth.
- Dynel, Marta. 2011. A Web of Deceit: A Neo-Gricean View on Types of Verbal Deception. *International Review of Pragmatics* 3 (2): 137-137. doi:10.1163/187731011X610996.
- Eco, Umberto. 1976. *Trattato di Semiotica Generale*. Milano: Bompiani. Translated as *A theory of Semiotics*. London: MacMillan.
- Egan, Andy & Weatherson, Brian (eds). 2011. *Epistemic Modality*. Oxford: OUP. DOI: [10.1093/acprof:oso/9780199591596.001.0001](https://doi.org/10.1093/acprof:oso/9780199591596.001.0001).
- Egan, Andy, Hawthorne, John & Weatherson, Brian. 2005. Epistemic modals in context. In *Contextualism in Philosophy: Knowledge, Meaning and Truth*, Gerhard Preyer & Georg Peter (eds). Oxford: Clarendon Press.
- Ekman, Paul. 1985. *Telling Lies: Clues to Deceit in the Marketplace, Marriage, and Politics*. New York: W.W. Norton.
- Engel, Pascal. 2007. Belief and normativity? *Disputatio*, 2 (23), 153-77
- Engel, Pascal. 2008. IN WHAT SENSE IS KNOWLEDGE THE NORM OF ASSERTION. *Grazer Philosophische Studien* 77 (1): 99-113.

- Engel, Pascal. 2013. In defence of normativism about the aim of belief. In Chan, T. (ed.) *The Aim of Belief*. Oxford: Oxford University Press.
- Eriksson, John. 2011. Straight Talk: Conceptions of Sincerity in Speech. *Philosophical Studies* 153 (2): 213–234. doi:10.1007/s11098-009-9487-2.
- Fadiman, Clifton. 1985. *The Little, Brown Book of Anecdotes*. Boston: Little, Brown.
- Falkenberg, Gabriel. 1988. Insincerity and disloyalty. *Argumentation* 2.1: 89-97.
- Falkenberg, Gabriel. 1990. Searle on sincerity. In *Speech Acts, Meaning and Intentions*. Walter de Gruyter, Berlin, 129-145.
- Fallis, Don. 2009. What Is Lying? *The Journal of Philosophy* 106 (1): 29–56.
- Fallis, Don. 2010. Lying and deception. *Philosophers' Imprint*, (10) 1–22.
- Fallis, Don. 2011. What liars can tell us about the knowledge norm of practical reasoning. *The Southern Journal of Philosophy* 49(4): 347–367. DOI: 10.1111/j.2041-6962.2011.00078.x.
- Fallis, Don. 2012. Lying as a Violation of Grice's First Maxim of Quality. *Dialectica* 66 (4): 563–581.
- Fallis, Don. 2013. Davidson was Almost Right about Lying. *Australasian Journal of Philosophy*. 91(2): 337–353. <http://doi.org/10.1080/00048402.2012.688980>
- Fallis, Don. 2015. Are Bald-Faced Lies Deceptive after All? *Ratio* 28(1):81-96. <http://doi.wiley.com/10.1111/rati.12055>.
- Faulkner, P., 2007. What is Wrong with Lying? *Philosophy and Phenomenological Research*, 75: 524–547.
- Faulkner, Paul. 2013. Lying and Deceit. In Hugh LaFollette (ed.), *International Encyclopedia of Ethics*. Hoboken, NJ: Wiley-Blackwell, 3101-3109.
- Feehan, Thomas D. 1988. Augustine on Lying and Deception. *Augustinian Studies*, 131–39.
- Feldman, R. S., Forrest, J. A., & Happ, B. R. 2002. Self-presentation and verbal deception: Do self-presenters lie more? *Basic and applied social psychology*, 24(2), 163-170.
- Foot, Philippa. 2001. *Natural Goodness*. Oxford: Oxford University Press.
- Frankfurt, H. G. 1999. The Faintest Passion. In *Necessity, Volition and Love*, Cambridge: Cambridge University Press, 95-107.
- Frankfurt, Harry G. 1986. On Bullshit. *Raritan Quarterly Review* 6:2, 81-100.
- Frankish, Keith. 2009. Partial belief and flat-out belief. In Franz Huber & Christoph Schmidt-Petri (eds), *Degrees of Belief* (75–93). Berlin: Springer. DOI: 10.1007/978-1-4020-9198-8_4.
- Fraser, Bruce. 1980. Conversational mitigation. *Journal of Pragmatics* 4: 341-350.
- Fraser, Bruce. 2010. Pragmatic Competence: The Case of Hedging. In *New Approaches to Hedging*, ed. G. Kaltenbo, W. Mihatsch & Schneider Stefan, 15–34. Emerald Group Publishing Limited.
- Frege, G. 1918. Der Gedanke. *Beiträge Zur Philosophie Des Deutschen Idealismus*. 58–77.
- Frege, Gottlob. 1879. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* [Concept Script, a formal language of pure thought modelled upon that of arithmetic]. Halle: L. Nebert. In *From Frege to Gödel: A Source Book in Mathematical Logic*, ed. by Jan van Heijenoort. Cambridge, MA: Harvard University Press.

- Frege, Gottlob. 1892. Über sinn und bedeutung, *Zeitschrift Für Philosophie Und Philosophische Kritik* [On sense and nominatum] 100: 22–50.
- Fried, Charles. 1978. *Right and Wrong*. Cambridge: Harvard
- Fried, Charles. 1981. *Contract as Promise*. Cambridge, MA: Harvard University Press.
- Friend, Stacie. 2008. Imagining Fact and Fiction. In *New Waves in Aesthetics*, edited by Kathleen Jones and Katherine Thomson-Jones. Palgrave Macmillan.
- Friend, Stacie. 2014. Believing in Stories. In *Aesthetics and the Sciences of Mind*, edited by Greg Currie, Matthew Kieran, Aaron Meskin, and Jon Robson, 227–48. doi:10.1093/acprof:oso/9780199669639.003.0012.
- Gale, Richard M. 1971. The Fictive Use of Language. *Philosophy* 46 (178) 324–40. doi:10.2307/3750012.
- García-Carpintero, Manuel. 2004. Assertion and the Semantics of Force-Markers. In C. Bianchi (ed.) *The Semantics/Pragmatics Distinction*: 133–166. Stanford: The University of Chicago Press.
- García-Carpintero, Manuel. 2013. Explicit Performatives Revisited. *Journal of Pragmatics* 49: 1–17.
- Gerken, M. 2014. Same, same but different: the epistemic norms of assertion, action and practical reasoning. *Philosophical Studies* 168(3): 725–744.
- Gerken, Mikkel. 2012. Discursive Justification and Skepticism. *Synthese* 189 (2): 373–94. doi:10.1007/s11229-012-0076-0.
- Ginet, Carl, 1979. Performativity. *Linguistics & Philosophy* 3: 245–265.
- Goldberg, S. 2015. *Assertion. On the philosophical significance of assertoric speech*. Oxford: Oxford University Press.
- Gottwald, Siegfried. 2001. *A treatise on many-valued logics*. Baldock: Research Studies Press.
- Graff, Gerald. 1980. *Poetic Statement and Critical Dogma*. Evanston: Northwestern University Press.
- Graff, Gerald. 1979. *Literature against Itself. Literary Ideas in Modern Society*. Chicago: The University of Chicago Press.
- Green, Mitchell S. 2000. Illocutionary Force and Semantic Content. *Linguistics and Philosophy* 23: 435–473.
- Green, Mitchell S. 2007. *Self-Expression*. Oxford: Oxford University Press.
- Green, Mitchell S. 2013. Assertions. In *Pragmatics of Speech Actions, Vol. II of the Handbook of Pragmatics*, II:0–33.
- Greenough, Patrick. 2011. Truth-Relativism, Norm-Relativism, and Assertion. In *Assertion: New Philosophical Essays*, 197–232. doi:10.1093/acprof.
- Grice, Herbert P. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Griffiths, P. J. 2004. *Lying: An Augustinian theology of duplicity*. Wipf and Stock Publishers.
- Grotius, Hugo. [RWP]. *The Rights of War and Peace*. F. W. Kelsey (trans.), Indianapolis: Bobbs-Merrill.
- Gussenhoven, Carlos. 2002. Intonation and interpretation: phonetics and phonology. In

- Proceedings of the Speech Prosody International Conference*, ed. by Stefan Sudhoff et al, 47-57. Aix en Provence: Laboratoire Parole et Langage.
- Hájek, Petr. 1998. *Metamathematics of fuzzy logic*. Dordrecht: Kluwer.
- Hardin, Karol J. forthcoming. Linguistic Approaches to Lying and Deception. In Meibauer, J. (ed.) *The Oxford Handbook on Lying*. Oxford: Oxford University Press.
- Hare, Richard. 1952. *The language of morals*. Oxford: Oxford University Press.
- Harris, Roy. 1978. The descriptive interpretation of performative utterances. *Journal of Linguistics* 14(2): 309-310.
- Hartman, N. 1975. *Ethics*. Translated by S. Coit. Atlantic Highlands, NJ: Humanities Press.
- Hawthorne, J. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.
- Heal, J. 1974. Explicit Performative Utterances and Statements. *The Philosophical Quarterly* 24(95). <http://doi.org/10.2307/2217715>
- Hedenius, Ingemar. 1963. Performatives. *Theoria* 29: 115-36.
- Hill, Christopher S, and Joshua Schechter. 2007. HAWTHORNE'S LOTTERY PUZZLE AND THE NATURE OF BELIEF. *Philosophical Issues* 17.
- Hinchman, Edward S. 2013. Assertion, Sincerity, and Knowledge. *Noûs* 47 (4): 613-46. doi:10.1111/nous.12045.
- Hindriks, F. 2007. The Status of the Knowledge Account of Assertion. *Linguistics and Philosophy* 30(3): 393- 406
- Hindriks, F. 2009. Constitutive Rules, Language, and Ontology. *Erkenntnis* 71 (2): 253-275.
- Hirst, Paul H. 1973. Literature and the Fine Arts as a Unique Form of Knowledge, *Cambridge Journal of Education* 3 (3): 118-132.
- Holmes, Janet. 1984. Modifying illocutionary force. *Journal of Pragmatics* 8: 345-365. DOI: 10.1016/0378-2166(84)90028-6.
- Holton, Richard. 2008. Partial Belief, Partial Intention. *Mind* 117: 27-58.
- Hoye, Leo. 1997. *Adverbs and modality in English*. London, New York: Longman.
- Huber, Franz & Christoph Schmidt-Petri (eds.). 2009. *Degrees of belief*. Springer.
- Humberstone, Lloyd. 1992. Direction of Fit. *Mind* 101 (401).
- Hume, David. [THN]. *A treatise on human nature*. NY: Oxford University Press.
- Isenberg, Arnold. 1964. Deontology and the Ethics of Lying. *Philosophy and Phenomenological Research* 24: 463-480.
- Jary, Mark. 2007. Are explicit performatives assertions? *Linguistics and Philosophy* 30(2): 207 - 234.
- Johnson, Casey Rebecca. 2017. What Norm of Assertion? *Acta Analytica*, May. Springer Netherlands, 1-17. doi:10.1007/s12136-017-0326-3.
- Juhl, P. D. 1980. *Interpretation. An Essay in the Philosophy of Literary Criticism*. Princeton: Princeton University Press
- Kant, Immanuel. [CPR]. *Critique of pure reason*. Norman Kemp Smith. London: Macmillan

- Kant, Immanuel. [GMM]. Groundwork of the Metaphysics of Morals. Trans. by M. J. Gregor (trans.). In A. W. Wood and M. J. Gregor (eds.), *Immanuel Kant, Practical Philosophy*. Cambridge: Cambridge University Press.
- Kant, Immanuel. [SRTL]. On a Supposed Right To Lie from Philanthropy. In Gregor, Mary J. (ed., trans.), *The Cambridge Edition of the Works of Immanuel Kant- Practical Philosophy*, 605-616. Cambridge: Cambridge University Press.
- Keiser, Jessica. 2015. Bald-Faced Lies: How to Make a Move in a Language Game without Making a Move in a Conversation. *Philosophical Studies* 1 (2014). Springer Netherlands: 251-64. doi:10.1007/s11098-015-0502-5.
- Kemp, Gary. 2007. Assertion as a Practice. In *Truth and Speech-Acts*, edited by D. Greimann and G. Siegart, 106-29. Routledge.
- Kenyon, Tim. 2003. CYNICAL ASSERTION: CONVENTION, PRAGMATICS, and SAYING 'UNCLE'. *American Philosophical Quarterly* 40 (3): 241-48.
- Kenyon, Tim. 2010. ASSERTION AND CAPITULATION. *Pacific Philosophical Quarterly* 91: 352-68.
- Kingsbury, Justine and Jonathan McKeown-Green. 2009. Definitions: Does Disjunction Mean Dysfunction? *Journal of Philosophy* 106:568-85
- Klein, Peter. 1998. Certainty. In Edward Craig, *Routledge Encyclopedia of Philosophy* 264-267.
- Knobe, Joshua. 2003. Intentional action and side effects in ordinary language. *Analysis* 279: 190-194.
- Koethe, John. 2009. Knowledge and the Norms of Assertion. *Australasian Journal of Philosophy* 87 (4): 625-38. doi:10.1080/00048400802598660.
- Kölbel, Max 2010. Literal Force: A Defence of Conventional Assertion. In Sarah Sawyer (ed.), *New Waves in Philosophy of Language*. London: Palgrave Macmillan, 108-37.
- Kratzer, Angelika. 1981. The notional category of modality. In Hans-Jürgen Eikmeyer & Hannes Rieser (eds), *Worlds, Words, and Contexts*, 38-74. Berlin: De Gruyter.
- Krishna, Daya. 1961. 'Lying' and the Compleat Robot. *The British Journal for the Philosophy of Science* 12 (46).
- Kupfer, Joseph. 1982. The moral presumption against lying. *The Review of Metaphysics* 36(1): 103-126.
- Kurthy, M., Lawford-Smith, H., & Sousa, P. (2017). Does ought imply can? *PloS one*, 12(4), e0175206.
- Kvanvig, J. 2009. Assertion, Knowledge, and Lotteries. In Greenough, P. & Pritchard, D. (eds.) *Williamson on Knowledge*. Oxford: Oxford University Press.
- Kvanvig, Jonathan L. 2011. Norms of Assertion. In J. Brown and H. Cappelen (eds.) *Assertion: New Philosophical Essays*. doi:10.1093/acprof.
- Labinaz, Paolo, and Marina Sbisà. 2014. Certainty and Uncertainty in Assertive Speech Acts. In *Communicating Certainty and Uncertainty in Medical, Supportive and Scientific Contexts*, Ilaria Riccioni, Carla Canestrari, Andrzej Zuczkowski, and Ramona Bongelli (eds.). Amsterdam: John Benjamins Publishing Company.
- Labov, W. 1984. Intensity. In *Georgetown University Round Table on Language and Linguistics*,

- ed. by Deborah Schiffrin 43-70. Washington: Georgetown University Press.
- Lackey, J. 2007. Norms of Assertion. *Nous* 41(4): 594-626.
- Lackey, Jennifer. 2013. Lies and Deception: An Unhappy Divorce. *Analysis* 73 (2): 236-48. doi:10.1093/analys/ant006.
- Lakoff, George. 1975. Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. *Journal of Philosophical Logic* 2 (4): 458-508.
- Lamarque, Peter & Stein Haugom Olsen. 1994. *Truth, Fiction, and Literature. A Philosophical Perspective*. Oxford: Clarendon Press
- Leland, Patrick R. 2015. Rational Responsibility and the Assertoric Character of Bald-Faced Lies. *Analysis* 75 (4): 550-554 doi:10.1093/analys/anv080.
- Lemmon, J.E. 1962. Sentences verifiable by their use. *Analysis* 12: 86-89.
- Leonard, Henry S. 1959. Interrogatives, Imperatives, Truth, Falsity and Lies. *Philosophy of Science* 26: 172-186.
- Levine, T. R. (ed.). 2014. *Encyclopedia of deception*. SAGE Publications.
- Levinson, Stephen. C. 1983. *Pragmatics*. Cambridge: CUP.
- Lewis, D. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8 (1):339-359.
- Lewis, David. 1970. General Semantics. *Synthese* 22: 18-67
- Lindley, T. Foster. 1971. Lying and Falsity. *Australasian Journal of Philosophy* 49: 152-157.
- Lyons, John, 1977. *Semantics*. Cambridge: Cambridge University Press
- MacCormick, N. (1983). WHAT IS WRONG WITH DECEIT? *Sydney Law Review*.
- MacFarlane, John. 2003. Epistemic Modalities and Relative Truth. Unpublished, Url=<<http://johnmacfarlane.net/epistmod-2003.pdf>>.
- MacFarlane, John. 2003. Future Contingents and Relative Truth. *The Philosophical Quarterly* 53 (212): 321-36. doi:10.1111/1467-9213.00315.
- MacFarlane, John. 2005. Making Sense of Relative Truth. *Proceedings of the Aristotelian Society*. 105 (1).
- MacFarlane, John. 2011. What is Assertion? In J. Brown and H. Cappelen (eds.) *Assertion: New Philosophical Essays*, 79-96. Oxford: Oxford University Press.
- MacIntyre, Alasdair. 1994. Truthfulness, Lies, and Moral Philosophers: What Can We Learn from Mill and Kant? *The Tanner Lectures on Human Values* 16: 307-61.
- Mahon, James Edwin. 2007. A Definition of Deceiving. *International Journal of Applied Philosophy* 21 (2).
- Mahon, James Edwin. 2008. Two Definitions of Lying. *International Journal of Applied Philosophy*, 211-230.
- Mahon, James Edwin. 2009. Why There are No Bald-Faced Lies. Paper presented at the Information Ethics Roundtable. <https://itunes.apple.com/it/podcast/why-there-are-no-bald-faced-lies/id413143120?i=1000092195610&mt=2>
- Mahon, James Edwin. 2010. The definition of lying, and why it matters. Roundtable at the School of Info Resources and Library Sciences, <https://itunesu.itunes.apple.com/feed/id413143120>.

- Mahon, James Edwin. 2011. Review of Lying and Deception by Thomas L. Carson, *Notre Dame Philosophical Reviews*. <http://ndpr.nd.edu/news/24572-lying-and-deception-theory-and-practice/>
- Mahon, James Edwin. 2016. The Definition of Lying and Deception, in Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2012 Edition).
- Maitra, Ishani. 2011. Assertion, Norms, and Games. In Brown, J. & Cappelen, H. (eds.). *Assertion: New Philosophical Essays*. Oxford: Oxford University Press.
- Mannison, Don S. 1969. Lying and lies. *Australasian Journal of Philosophy* 2(47): 132-144. DOI: 10.1080/00048406912341141.
- Margolis, Joseph 1965, *The Language of Art and Art Criticism*. Detroit: Wayne State University Press.
- Margolis, Joseph. 1962. "Lying Is Wrong" and "Lying Is Not Always Wrong". *Philosophy and Phenomenological Research* 23: 414-418.
- Marsili, Neri. 2014. Lying as a Scalar Phenomenon: Insincerity along the Certainty-Uncertainty Continuum. In S. Cantarini, W. Abraham, and E. Leiss (eds.). *Certainty-Uncertainty - and the Attitudinal Space in between*, 153-173. Amsterdam: John Benjamins Publishing Company. [10.1075/slcs.165.09mar](https://doi.org/10.1075/slcs.165.09mar)
- Marsili, Neri. 2015. Normative accounts of assertion: from Peirce to Williamson, and back again. *Rivista Italiana di Filosofia del Linguaggio* (2).
- Marsili, Neri. 2016. Lying by Promising. *International Review of Pragmatics*, 8(2), 271-313. <http://doi.org/10.1163/18773109-00802005>
- Marsili, Neri. 2017. Lying and Certainty. In J. Meibauer, *The Oxford Handbook on Lying*. Oxford: Oxford University Press.
- Marušić, Berislav. 2012. Belief and difficult action. *Philosopher's Imprint* 12(18).
- Marušić, Berislav. 2013. Promising against the Evidence. *Ethics* 123(2): 292-317.
- McCammon, C. 2014. Representing Yourself as Knowing. *American Philosophical Quarterly* 51(2): 1-14.
- McKenna, Robin. 2015. Assertion, Complexity, and Sincerity. *Australasian Journal of Philosophy* 93 (4) (October 2): 782-798.
- Meibauer, Jörg . 2005. Lying and Falsely Implicating. *Journal of Pragmatics* 37(9): 1373-1399.
- Meibauer, Jörg. 2011. On Lying: Intentionality, Implicature, and Imprecision. *Intercultural Pragmatics* 2 (8): 277-292.
- Meibauer, Jörg. 2014. *Lying at the semantic-pragmatic interface*. Berlin: De Gruyter.
- Meibauer, Jörg. 2014b. Bald-Faced Lies as Acts of Verbal Aggression. *Journal of Language Aggression and Conflict* 2 (1): 127-50. doi:10.1075/jlac.2.1.05mei.
- Meibauer, Jörg. 2014c. A Truth That's Told with Bad Intent: Lying and Implicit Content. *Belgian Journal of Linguistics* 28: 97-118. doi:10.1075/bjl.28.05mei.
- Meibauer, Jörg. 2016. Understanding Bald-Faced Lies. *International Review of Pragmatics* 8 (2): 247-70. doi:10.1163/18773109-00802004.
- Mele, Alfred. 1992. *Springs of Action*. Oxford: Oxford University Press.
- Mikhail, J. 2011. *Elements of moral cognition: Rawls' linguistic analogy and the cognitive science*

- of moral and legal judgment*. Cambridge University Press.
- Mikkonen, Jukka. 2010. Implicit Assertions in Literary Fiction. *Proceedings of the European Society for Aesthetics* 2 (1960): 312–30.
- Montaigne, Michel. 1595 [E]. *Essais*. Verdun: P. Villey et Saulnier.
- Montminy, M. 2013. The Single Norm of Assertion. In A. Capone, F. Lo Piparo, & M. Carapezza (Eds.), *Perspectives on Pragmatics and Philosophy*: 35–52.
- Moore, George. 1966. *Ethics* (2nd ed.), Oxford: Oxford University Press.
- Moore, George. 1993. *Selected Essays*. London: Routledge.
- Moran, Richard. 2005. Getting Told and Being Believed. *Philosophers' Imprint* 5(5).
- Moran, Richard. 2005. Problems with sincerity. *Proceedings of the Aristotelian Society* 105: 325–345.
- Nagel, T., 1970. *The Possibility of Altruism*. Princeton: Princeton University Press.
- Newey, G., 1997. Political Lying: A Defense. *Public Affairs Quarterly* 11: 93–116.
- Owens, David. 2008. Promising without Intending. *The Journal of Philosophy* 105(12): 737–755.
- Pagin, Peter. 2015. Assertion, In E. N. Zalta (ed.), *Stanford Encyclopaedia of Philosophy (Spring 2015 edition)*, <<http://plato.stanford.edu/archives/spr2015/entries/assertion/>>.
- Pagin, Peter. 2016. Problems with Norms of Assertion. *Philosophy and Phenomenological Research* 93 (1): 178–207. doi:10.1111/phpr.12209.
- Papafragou, A. 2006. Epistemic modality and truth conditions. *Lingua* 116(10), 1688–1702. <http://doi.org/10.1016/j.lingua.2005.05.009>.
- Parsons, T. 1978. Review of John Woods: Logic of Fiction. *Synthese*, 39, 155–164.
- Peirce, Charles S. [CP] *Collected Papers of Charles Sanders Peirce*, 8 vols. Edited by Charles Hartshorne, Paul Weiss, and Arthur W. Burks (Harvard University Press, Cambridge, Massachusetts, 1931–1958).
- Peirce, Charles S. [MS] *The Charles S. Peirce Papers* (Cambridge: Harvard University Library, 1966, microfilm, 33 reels including supplement)
- Pelling, Charlie. 2013. Assertion and Safety. *Synthese* 190 (17): 3777–96. doi:10.1007/s11229-012-0223-7.
- Pelling, C. 2013b. Assertion and the Provision of Knowledge. *The Philosophical Quarterly* 63(251). <http://doi.org/10.1111/1467-9213.12013>
- Pepp, Jessica. forthcoming. *Truth Serum, Liar Serum, and Some Problems about Saying what You Think is False*. In E. Michaelson and A. Stokke (eds.), *Lying*. Oxford: Oxford University Press.
- Peter Lombard. [SEN] *Libri Quattuor Sententiarum*.
- Plantinga, Alvin. 1978. *The nature of necessity*. Oxford: Oxford University Press.
- Plunze, Christian. 2001. Try to Make Your Contribution One that is True. *Acta Philosophica Fennica* 69:177–89.
- Poggi, I., D'Errico, F., & Vincze, L. 2011. Discrediting moves in political debates. In *Proceedings of second international workshop on user models for motivational systems: the affective and the rational routes to persuasion*. Springer LNCS: 84–99.

- Pollock, JL. 1982. *Language and Thought*. Princeton: Princeton University Press
- Price, Huw. 1983. Does 'Probably' Modify Sense? *Australasian Journal of Philosophy*: 37-41.
- Primoratz, Igor. 1984. Lying and the 'Methods of Ethics', *International Studies in Philosophy* XVI:35-57
- Pritchard, D. 2013. Epistemic luck, safety, and assertion. In J. Turri & C. Littlejohn (Eds.), *Epistemic Norms: New Essays on Action, Belief, and Assertion*. Oxford: Oxford University Press.
- Rawls, J. 1955. Two Concepts of Rules. *The Philosophical Review* 64(1): 3-32.
- Rawls, John. 1981. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Récanati, François. 1987. *Meaning and Force: The Pragmatics of Performative Utterances*. Cambridge University Press, Cambridge
- Reed, Baron. 2008. Certainty. In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL=<http://plato.stanford.edu/archives/win2011/entries/certainty>.
- Reichert, John. 1981. Do Poets Ever Mean What They Say? *New Literary History*. 13 (1):53-68.
- Reichert, John. 1977. *Making Sense of Literature*. Chicago: University of Chicago Press.
- Reimer, Marga. 1995. Performative utterances: a reply to Bach and Harnish. *Linguistics and Philosophy* 18: 655-675.
- Reimer, Marga. 2004. What malapropisms mean: A reply to Donald Davidson, *Erkenntnis* 60(3): 317- 334.
- Rescorla, M. 2007. A Linguistic Reason for Truthfulness. In *Truth and Speech Acts*, edited by Dirk Greimann and Geo Siegwart, Routledge.
- Rescorla, M. 2009. Assertion and Its Constitutive Norms. *Philosophy and Phenomenological Research* LXXIX (1): 98-130.
- Reynolds, Steven L. 2002. Testimony, Knowledge, and Epistemic Goals. *Philosophical Studies* 110 (2): 139-61. doi:10.1023/A:1020254327114.
- Richard Ohmann. 1971. Speech Acts and the Definition of Literature. *Philosophy & Rhetoric* 4 (1).
<https://www.jstor.org/stable/pdf/40236740.pdf?refreqid=excelsior%3A59add992bef1d42deff9f8a31a4df7b>.
- Ridge, Michael. 2006. Sincerity and expressivism. *Philosophical Studies* 131: 487-510.
<http://doi.org/10.1007/s>
- Ridge, Michael. 2011. Reasons for Action: Agent-Neutral vs. Agent-Relative. In E. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition) URL = <https://plato.stanford.edu/archives/win2011/entries/reasons-agent/>.
- Ross, David. 1930. *The Right and the Good*, reprinted in 2002. Oxford: Oxford University Press.
- Rowe, M. W. 1997. Lamarque and Olsen on Literature and Truth. *The Philosophical Quarterly* 47(188): 322-41.
- Rutschmann, Ronja, and Alex Wiegmann. 2017. No Need for an Intention to Deceive? Challenging the Traditional Definition of Lying. *Philosophical Psychology*. doi:10.1080/09515089.2016.1277382.
- Ryan S. 2003. Doxastic compatibilism and the ethics of belief. *Philosophical Studies*. 114(1-2):47-

- Saul, Jennifer M. 2011. Just go ahead and lie. *Analysis* 72(1): 3-9.
- Saul, Jennifer M. 2012. *Lying, Misleading, and what is Said: An Exploration in Philosophy of Language and in Ethics*. Oxford: Oxford University Press.
- Sbisà, Marina. 2001. Illocutionary force and degrees of strength in language use. *Journal of Pragmatics* 33: 1791-1814. DOI: 10.1016/S0378-2166(00)00060-6
- Sbisà, Marina. 2016. Varieties of Speech Act Norms. In *Dynamics and Varieties of Speech Actions*. doi:10.1017/CBO9781107415324.004.
- Schaffer, J. 2008. Knowledge in the image of assertion. *Philosophical issues* 18(1).
- Schopenhauer, Arthur. 1974. *Parerga e Paralipomena*. Translated by EFJ Payne (vol 2).
- Schwyzler, H. 196). Rules and Practices. *Philosophical Review* 78(4).
- Searle, John R and Daniel Vanderveken. 1985. *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.
- Searle, John R. 1964. How to Derive Ought from Is. *The Philosophical Review* 73(1), 43-58
- Searle, John R. 1965. What is a speech act? In Black, M.(ed.) *Philosophy in America*, 221-39. London, Allen & Unwin.
- Searle, John R. 1969. *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139173438
- Searle, John R. 1975. The Logical Status of Fictional Discourse. *New Literary History* 6 (2): 319-32.
- Searle, John R. 1976. A Classification of Illocutionary Acts. *Language in Society* 5 (1): 1-23.
- Searle, John R. 1989. How performatives work. *Linguistics & Philosophy* 12:535-558.
- Searle, John R. 1995. *The Construction of Social Reality*. New York, The Free Press.
- Searle, John R. 2007. Illocutionary acts and the concept of truth. In *Truth and Speech-Acts*, edited by D. Greimann and G. Siegart, 106-29. Routledge.
- Searle, John R., and Daniel Vanderveken. 2005. Speech acts and illocutionary logic. *Logic, Thought and Action*. Springer Netherlands. 109-132.
- Seierstad, Asne. 2003. *A Hundred and One Days: A Baghdad Journal*, trans. by Ingrid Christophersen. New York: Basic Books.
- Shah, Nishi .2003. How truth governs belief . *Philosophical Review* 112: 447-82 .
- Shah, Nishi and J. David Velleman. 2005. Doxastic Deliberation. *Philosophical Review* 114: 497-534.
- Sidgwick, Henry. 1981 [1874]. *The Methods of Ethics*. Indianapolis. IN: Hackett.
- Sidney, S. P. 1595. *An Apology For Poetry (Or The Defence Of Poesy)*. Manchester: Manchester University Press (Revised and Expanded Second Edition, 2002).
- Siegler, Frederick A. 1966. Lying. *American Philosophical Quarterly* 3: 128-136.
- Simpson, David. 1992. Lying, Liars and Language, in *Philosophy and Phenomenological Research* 52: 623-639.

- Smith, David. L. 2004. *Why We Lie: The Evolutionary Roots of Deception and the Unconscious Mind*. St. Martin's Press.
- Sorensen, Roy. 2007. Bald Faced Lies! Lying without the intent to deceive. *Pacific Philosophical Quarterly* 88: 251-264.
- Sorensen, Roy. 2010. Knowledge-Lies. *Analysis* 70 (4) (August 7): 608-615. doi:10.1093/analys/anq072.
- Sorensen, Roy. 2011. What lies behind misspeaking. *American Philosophical Quarterly* 48(4): 399-409.
- Sosa, David. 2009. Dubious Assertions. *Philosophical Studies* 146 (August 2008): 269-72. doi:10.1007/s11098-008-9255-8.
- Staffel, Julia. 2011. Reply to Sorensen, 'knowledge-lies'. *Analysis* 71: 300-303.
- Staffel, Julia. 2012. Can There Be Reasoning with Degrees of Belief? *Synthese* 2011: 1-20
- Stalnaker, Robert. 1978. Assertion. In R. Stalnaker, *Context and content*, 78-95. Oxford: Oxford University Press.
- Stalnaker, Robert. 1998. On the representation of context. In R. Stalnaker, *Context and content*, 96-114. Oxford: Oxford University Press.
- Stanley, Jason. 2008. Knowledge and certainty. *Philosophical Issues* 18.1 35-5.
- Stern R. 2004. Does 'ought' imply 'can'? And did Kant think it does? *Utilitas*.16(1):42-61.
- Stokke, Andreas. 2013. Lying and asserting. *Journal of Philosophy* 110(1): 33-60.
- Stokke, Andreas. 2013b. Lying, Deceiving, and Misleading. *Philosophy Compass* 8: 348-359.
- Stokke, Andreas. 2014. Insincerity. *Noûs* 48 (3): 496-520. doi:10.1111/nous.12001.
- Stokke, Andreas. 2016. Proposing, Pretending, and Propriety: A Response to Don Fallis. *Australasian Journal of Philosophy* (May): 1-6. doi:10.1080/00048402.2016.1185739.
- Stone, Jim. 2007. Contextualism and Warranted Assertion. *Pacific Philosophical Quarterly* 88: 92-113. doi:doi:10.1111/j.1468-0114.2007.00282.x.
- Swanson, Eric. 2011. How not to theorize about the language of subjective uncertainty. In *Epistemic Modality*, Andy Egan & Brian Weatherson (eds), 249-269. Oxford: OUP. DOI: 10.1093/acprof:oso/9780199591596.003.0009.
- Sweetser, Eve. 1987. The definition of lie: an examination of the folk models underlying a semantic prototype, in D. Holland and Q. Naomi (eds.), *Cultural Models in Language and Thought*, 3-66. Cambridge: Cambridge University Press.
- Swift, Jonathan. 1710. The Art of Political Lying. *The Examiner*, 10.
- Turri, Angelo and John Turri. 2015. The Truth about Lying. *Cognition*. 138: 161-168.
- Turri, J. 2013. The Test of Truth: An Experimental Investigation of the Norm of Assertion. *Cognition* 129 (2): 279-291.
- Turri, John, and Peter Blouw. 2015. "Excuse Validation: A Study in Rule-Breaking." *Philosophical Studies* 172 (3): 615-34.
- Turri, John. 2010. Epistemic Invariantism and Speech Act Contextualism. *Philosophical Review* 119 (1): 77-95.

- Turri, John. 2011. The Express Knowledge Account of Assertion. *Australasian Journal of Philosophy* 89 (1): 37-45.
- Turri, John. 2013. Knowledge and Suberogatory Assertion. *Philosophical Studies* 167 (3): 557-67.
- Urmson, J. O. 1976. *Fiction* American Philosophical Quarterly 13 (2): 153-157.
- Van der Henst, Jean-Baptiste, Laure Carles & Dan Sperber. 2002. Truthfulness and Relevance in Telling The Time. *Mind and Language*, 17 (5) (November): 457-466.
- Van der Schaar, Maria. 2011. Assertion and Grounding: A Theory of Assertion for Constructive Type Theory. *Synthese* 183: 187-210. doi:10.1007/s11229-010-9758-7.
- Van Inwagen, Peter. 1977. Creatures of Fiction. *American Philosophical Quarterly* 14 (4): 299-308.
- Vanderveken, D. 1990. *Meaning and Speech Acts: Volume 1, Principles of Language Use* (Vol. 1). Cambridge University Press.
- Vanderveken, Daniel. 1980. Illocutionary logic and self-defeating speech acts. In J.R. Searle, F. Kiefer and M. Bierwisch, *Speech act theory and pragmatics*, 247-272. Amsterdam: Springer.
- Viebahn, E. 2017. *Non-literal Lies*. Erkenntnis. <http://doi.org/10.1007/s10670-017-9880-8>
- Vlach, F. 1981. Speaker's meaning. *Linguistics and Philosophy*. 4 (3),359-391.
- Von Fintel, Kai, & Anthony S. Gillies. 2008. CIA leaks. *Philosophical review*, 117(1), 77-98. doi: 10.1215/00318108-2007-025
- Vrij, A. 2008. *Detecting Lies and Deceit. Pitfalls and Opportunities*. 2nd ed, 1st ed. 2000. Chichester: Wiley
- Walton, Kendall. 1990. *Mimesis as make-believe: On the foundations of the representational arts*. Harvard University Press.
- Watson, Gary. 2004. Asserting and Promising. *Philosophical Studies* 117(1): 57-77.
- Wedgwood, Ralph. 2002. The Aim of Belief. *Philosophical Perspectives*, 16, 267-97.
- Weiner, Matthew. 2005. Must We Know What We Say? *Philosophical Review* 114 (2): 227-251.
- Whiting, Daniel. 2010. Should I Believe the Truth? *Dialectica* 64 (2): 213-24. doi:10.1111/j.1746-8361.2009.01204.x.
- Whiting, Daniel. 2012. Stick to the Facts: On the Norms of Assertion. *Erkenntnis* 78 (4): 847-867.
- Whiting, Daniel. 2013. The Good and the True (or the Bad and the False). *Philosophy* 88: 219-242
- Wiegmann, Alex, Jana Samland & Michael R. Waldmann, 2016. Lying despite telling the truth. *Cognition* 150: 37-42.
- Williams, Bernard. 1966. Consistency and realism, *Proceedings of the Aristotelian Society*, 40: 1-22.
- Williams, Bernard. 1985. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.
- Williams, Bernard. 2002. *Truth and Truthfulness: An Essay in Genealogy*. Princeton NJ: Princeton University Press .
- Williams, John N. 1996. Moorean Absurdities and the Nature of Assertion. *Australasian Journal*

- of Philosophy* 74 (September 2014): 135–49. doi:10.1080/00048409612347111.
- Williamson, Timothy. 1996. Knowing and Asserting. *The Philosophical Review* 105, 4, 489–523.
- Williamson, Timothy. 2000. Assertion. In *Knowledge and Its Limits*, 238–270. Oxford University Press.
- Wilson, Deirdre & Dan Sperber. 2002. Truthfulness and Relevance. *Mind* 25: 1–41.
- Wittgenstein, Ludwig. [PI]. *Philosophical investigations*. 4th ed., trans. by Anscombe, Hacker & Schulte. Wiley-Blackwell
- Wood, D., 1973. Honesty. In A. Montefiore (ed.), *Philosophy and Personal Relations: An Anglo-French Study*, London: Routledge, 192–218.
- Wright, C. 1992. *Truth and Objectivity*. Cambridge, Mass.: Harvard University Press.
- Xu, Fen, Yang C. Luo, Genyue Fu, and Kang Lee. 2009. Children’s and adults’ conceptualization and evaluation of lying and truth-telling. *Infant and Child Development* 18(4): 307–322.
- Yalcin, Seth. 2007. Epistemic Modals. *Mind*, 116(464), 983–1026. <http://doi.org/10.1093/mind/fzm983>
- Yalcin, Seth. 2011. Nonfactualism about epistemic modality. In Andy Egan & B. Weatherson (eds.), *Epistemic Modality*. Oxford University Press.