

Coercion and Lies

Katherine Hawley, version of January 2017

Can we be coerced into lying? Or does the very fact of coercion undercut the possibility of making an assertion? Through discussion of capitulations and other forms of coerced speech, this chapter explores the ways in which apparent assertions may be drained of standard normative significance, and thus excluded from the category of lies. Coerced pseudo-assertions are in this way similar to coerced pseudo-promises, and to coerced pseudo-gifts, neither of which have the standard normative significance associated with genuine promises and gifts. Nevertheless, our speech and actions under coercion are liable to moral evaluation, and coercion does not always make it permissible to speak falsely or attempt to mislead an audience.

Keywords: coercion, duress, lies, assertion, promises, gifts.

Paradigmatic assertions are freely-chosen actions, and paradigmatic lies are freely-chosen perversions of this practice. But we are often under external pressure to speak, or to remain silent, or to say something in particular. Such pressure can range from slight social embarrassment to physical or mental torture, and from passing temptation to irresistible bribes. Some pressure is obviously coercive, whilst some is obviously not; the intermediate third category is large and confusing.

There is a wealth of legal scholarship concerning compelled speech and its relationship to freedom of speech principles (e.g. Brudney 2000), the admissibility and reliability of confessions and testimony provided under duress (e.g. Perillo and Kassin 2011), and the ways in which duress, undue influence, or indeed deception can invalidate contracts or consents (e.g. Chen-Wishart 2015). Moreover philosophers have written extensively about coerced action in general, attempting both to define coercion, and to understand its moral significance for both coercer and coercee: when and how does coercion excuse or justify behaviour which would otherwise be impermissible? (Anderson 2015)

But coerced speech, or speech elicited under duress, is rarely the main focus of attention in the philosophical literature. One key exception is debate over whether apparent promises offered under coercion can be binding (e.g. Shiffrin 2014), and I will return to this topic

below. Another is exciting new work by Rachel McKinney, in which she develops the notion of ‘extracted speech’, paying particular attention to what she calls ‘[u]njust extracted speech...[i.e.] speech that an agent is made to produce that also wrongs her’ (2016: 259). McKinney focuses on power structures and subordination; my approach and focus are somewhat different to McKinney’s, but I believe our discussions are usefully complementary.

In this paper I address a range of possible situations in which it seems that someone is coerced into telling a lie – or, at least, coerced into doing something which would have counted as lying had it not been coerced. It is clear that people can be coerced into saying things which they know to be false, and indeed coerced into attempting to deceive others, sometimes successfully. But can coerced utterances of this kind count as lies, or does coercion somehow nullify this possibility, just as coercion may sometimes invalidate what would otherwise have been a binding promise? I will also discuss the extent to which we can morally evaluate such coerced utterances even if they do not qualify as lies: coercion is widely thought to mitigate responsibility somehow, and this has consequences for coerced speech just as it has for coerced action.

Section 1: Capitulations

To motivate these discussions, I begin with an exchange between Tim Kenyon (2003, 2010) and Roy Sorensen (2007). Kenyon and Sorensen focus primarily on uncoerced yet obviously insincere utterances, where the speaker does not intend to deceive the audience – so-called ‘bald-faced lies’. They agree that such utterances can be assertions, but disagree about whether they qualify as lies (see also Maitra, this volume).

But they also discuss coerced utterances, and the ‘capitulations’ which sometimes occur when someone is under severe pressure to produce a specific utterance. For example, a victim may be tortured or threatened into renouncing religious or ideological convictions, or into confessing to some crime which she may or may not have committed. Both Kenyon and Sorensen cite Dummett’s example of a Christian subject to religious persecution, who is forced to trample on a crucifix; such episodes of persecution are central to Shūsako Endō’s 1966 novel *Silence*.

Dummett writes:

...in the same spirit, what [the Christian] is forced to do may be to *say*, e.g., “Jesus Christ is the offspring of the devil”... The victim may know that his persecutors will be quite aware that, even if he says what they want him to, he will not believe it: what is important to both of them is whether he *says* it or not. (Dummett 1981: 331, italics in original).

The persecuted man does not intend that his audience believe that his speech is an expression of his belief; in that sense, the utterance is patently insincere, as in a bald-faced lie. But for Kenyon, ‘To deny one’s faith under torture is not to assert’ (2003: 245); Leland (2015) makes a similar point. Sorensen (2007) disagrees, on several grounds. First, he argues that it is difficult to understand why victims are so desperate to avoid voicing capitulations if these are not really assertions. Second, he cites Gilbert (1993) in support of the claim that a coerced promise can create an obligation. Third, he argues that much speech is subject to some degree of duress, but is nevertheless described as ‘assertion’. Sorensen’s view seems to be that a coerced capitulation can be a (forgivable) lie.

In response, Kenyon (2010) compares capitulation to theatrical speech, which is uncontroversially non-assertoric. Kenyon argues that an audience which knows that a capitulation has been forced should not hold the speaker to any supposed ‘commitments’ of defence-or-retraction, and is not licensed to act on the basis of what is said. He explains the horror of capitulation not in terms of forced *assertion*, as Sorensen does, but in terms of ‘pride, dignity, and value for self-determination’ (2010: 364).

Religious cases have special complications, with implications of blasphemy and a divine audience. A different context of capitulation is the totalitarian state, in which people are perpetually uncertain of who is listening, and where their loyalties lie, with even family members under suspicion. Sorensen discusses Iraq under Saddam Hussein: ordinary people must constantly praise the dictator, to one another and to foreign visitors. For Sorensen, these utterances are assertions (and lies) despite everyone’s mutual awareness of the pragmatic reasons for speaking this way, and their lack of intention to deceive one another. These things must simply be said, for fear of the consequences otherwise.

Jennifer Saul, in discussing such cases, writes:

It seems somewhat reasonable to suggest that, since everyone is forced to make these false utterances, and everyone knows they are false, they cease to be genuine lies. The reason for this might even be that they cease to have their apparent content [as idioms do]...Perhaps they are even treated as meaningless. I think this story is possible, but that something is lost by adopting it wholesale. What is lost is a recognition of the fact that one of the terrible things done by a totalitarian state is to *force people to lie* with these state-supporting utterances...On the other hand, treating the utterances as meaningless might be an important source of resistance. (2012: 9-10)

These cases are incidental to Saul's main discussion. But it is striking that she focuses on content rather than force – either such capitulations are lies or else they lack (literal) content – thus omitting the possibility that such utterances are like theatrical performances, which have meaning but lack assertoric force. (Saul writes of warranting rather than assertion, but this serves to distinguish serious from non-serious contexts much as talk of assertion does.) Kenyon might respond that one awful aspect of living under such a regime is the constant need to perform as if appearing in a drearily predictable play.

Capitulations of the types discussed in this literature differ from ordinary assertions, even from ordinary lying assertions, in at least two ways. First, they embody a bald-faced lack of sincerity. The audience is aware that the speaker's immediate motive in saying that *p* is not to express a belief that *p*, or to persuade others that *p*, but simply to satisfy the tormenter's demands that she say that *p*. Indeed, the speaker is aware that the audience is aware of this, and so on, so there is no intent to deceive. It is this lack of sincerity which most concerns Kenyon, Sorensen, and Saul.

Second, the speaker is subject to coercion: this may affect the normative status of the speaker, the normative status of her speech, or indeed the normative status of the coercer. What about the audience? In some cases, coercer and audience are identified: we easily imagine Dummett's religious persecution case in this way. In other cases, the audience includes morally innocent third parties, perhaps including fellow victims of coercion: this is the story of everyday life in the totalitarian state. Morally innocent audiences can be further divided, into audiences who are aware that the speaker is coerced (e.g. fellow subjects of the

totalitarian state) and audiences who are not aware that the speaker is coerced (e.g. naïve visitors to the totalitarian state).

These two features – obvious lack of sincerity, and coercion – can come apart. We have the first feature without the second in standard cases of uncoerced bald-faced ‘lies’, e.g. the student who denies plagiarism for procedural reasons, although it is obvious to everyone concerned that she is guilty (Carson 2006). And we have the second feature without the first in several different types of situation. For example, a speaker may be coerced into saying that *p* to some audience, and moreover coerced into appearing sincere whilst she does so. Alternatively, a speaker may be coerced with the aim of getting her to reveal what she believes about whether *p*, i.e. to say whichever of *p* or *not-p* she believes, rather than coerced to say that *p* as such. Here, even if the audience consists of the coercers themselves, the speaker may appear sincere in saying whether *p*, and may in fact be genuinely sincere. In what follows, I will consider cases of coerced speech both with and without apparent sincerity.

Typically we think of coercion as making us do things we would not have done if uncoerced. But what if someone is coerced to say that *p*, even though she would have said that *p* even without coercion? For example, imagine that the totalitarian state requires everyone to say how beautiful the native landscape is, and yet many subjects genuinely believe that the landscape is beautiful, and would spontaneously say this without coercion. This is a point at which different accounts of coercion, and different accounts of responsibility and freedom, may issue different verdicts; we need carefully to distinguish counterfactuals from both causal and reason-giving claims in order to understand the relationship between coercion and action in such cases. To avoid such complications, I will focus on situations in which coercion is directly responsible for our saying and doing things we would not have said and done otherwise.

I am making distinctions between types of case, and I will make further distinctions below. Distinction-drawing is an ingrained philosophical habit, and for good reason: it is often a useful first step in attempting to understand complex issues. But I acknowledge that many real life cases will not easily be classifiable in terms of these distinctions. For example, the boundary between coercive interrogation (coercion to say sincerely whether *p* or *not-p*) and coerced capitulation (coercion to say that *p*) is often blurry. An interrogator may think he

already knows that the target believes that *p*, rather than *not-p* – for example the interrogator may attempt to coerce what he takes to be a sincere confession of a crime. (In this context, Kenyon (2010) references Henry Shue’s (1978) distinction between interrogational and terroristic torture.)

Moreover, there are related but different distinctions I might have drawn. For example, rather than distinguishing coercion to say sincerely whether *p* from coercion to say that *p*, I might have distinguished information-seeking from non-information-seeking coercion of speech. That is, I might have taxonomised cases in terms of what the coercer hopes to gain from the coercion, rather than in terms of what the coercee is expected to do as a result of the coercion. This would be important in a different paper, one more focused on the epistemology of coerced testimony than on the ethics of speaking under coercion. What I offer in this paper is thus not intended to be a full or final taxonomy of coerced speech: different categories will be useful for different local purposes, as for example McKinney’s (2016) framework works well for the questions she pursues.

Terrorised capitulations and coerced confessions are evidently atypical assertions and lies, if they are assertions and lies at all. As such, it makes sense to allow our more general accounts of assertion and lying to help determine how best to classify such cases, rather than relying solely upon these cases as crucial experiments which can fix which account of assertion or lying we should adopt. Nevertheless, thinking about coerced speech may help us better understand the ways in which we more standardly hold people responsible for what they say, including the lies they tell. After all, coercion is commonly thought to eliminate or at least mitigate the victim’s responsibility for her actions.

Section 2. Coercion and Responsibility

How might coercion affect the moral or other normative status of speech, especially knowingly-false speech? Unfortunately, there is no philosophical consensus about what constitutes coercion, nor how coercion affects the victim’s responsibility for what she is coerced to do (Anderson 2015 provides an overview of debate). Can there be coercive offers as well as coercive threats? What is the difference between a threat, an offer (-you-can’t-refuse), and even a warning? How threatening must a threat be in order to be coercive? Does coercion always wrong its target? I will not attempt to delineate borderline cases, nor to

define coercion. Rather, I will focus on cases in which it is clear that coercion takes place, in order to examine the consequences of such coercion.

One thing which does seem relatively uncontroversial is that we should distinguish between coercion which completely overwhelms the target's will, making her incapable of choosing how to act in the relevant matter, and coercion which changes the relative appeal of the options between which the target is choosing. This distinction applies both to speech and to non-speech forms of action.

Physical imprisonment coerces a target in the first way, rendering her unable to choose whether to leave the room or not. Threatening to kill the target's children if she leaves the room would typically coerce in the second way: leaving the room becomes highly unappealing to the target, but she is still in some sense able to choose whether to do so. Although physical imprisonment or force is the paradigm case of overwhelming coercion, the important distinction here is not between two types of coercive action, but between two ways in which coercion can affect its target. After all, the very same type of threat might coerce different people in different ways: some people might become physiologically paralysed upon hearing the threat to their children's lives, rendered unable to choose whether to leave the room.

How does this distinction apply to speech? A literal gag constitutes overwhelming coercion to literal silence. If we think more broadly of speech as including the ability to reach an audience, then overwhelming coercion to 'silence' could include solitary confinement, internet access restrictions, or confiscation of broadcasting equipment. There are obvious connections here to debate about what is protected by any right to free speech, and to debate about pornography and silencing (e.g. Maitra and McGowan 2012).

Can one be literally forced to speak, rather than remain silent? Can one be literally forced to say something in particular? Maybe hypnosis sometimes functions like this, or torture which entirely overwhelms the will, perhaps through mental dislocation and disorientation; forced speech might also result from an effective 'truth serum' if such a thing existed.

Overwhelming coercion is easier to implement for some types of communication than for others: imagine a hand physically forced to write, or a BSL/ASL signer with limbs literally strung like a puppet. We can distinguish overwhelming coercion to say that *p*, reducing us to

the level of a ventriloquist's dummy, from overwhelming coercion to expose one's belief as to whether *p*, making us readable as by a fantasy MRI scan which was sensitive to semantic content.

I will take it that someone who speaks as the result of such overwhelming coercion is not making an assertion, or warranting. This is because such episodes of speech are not intentional actions on the part of the speaker; likewise, sleep-talk doesn't count as assertion. This point does not rely upon any distinctive feature of speech as opposed to other behaviour: similarly, someone hypnotised into firing a gun is not performing an intentional action, and nor is someone physically forced to pull the trigger. Even without considering coercion, there are tricky borderline cases of assertion, including cases in which, perhaps because of linguistic misapprehension, someone might be intentionally asserting, yet unintentionally asserting that *p* (Alston 2000: section 5.8), or cases in which an utterance is a borderline assertion, perhaps because the speaker is half-asleep, or because she is wavering between asserting and merely suggesting. Nevertheless, it seems clear that speech, writing, or signing which is extracted through overwhelming coercion cannot constitute assertion.

There is a sense in which forced speech of this kind involves telling an audience something. But this is the 'natural meaning' sense of telling (Grice 1957), i.e. the sense in which the car in my driveway tells the neighbours that I am home, or the bags under my eyes tell my friends that I am not sleeping well. What someone says when overwhelmingly coerced may in this sense tell the audience something about what she believes, or indeed about how easy or difficult it is to overwhelm her will. But the speaker herself is not telling the audience that *p* even when she says that *p* under such circumstances, any more than she would tell an audience that *p* if someone forced her hand to write '*p*' (mysteriously!) on the page.

I will return to telling, asserting and warranting below when I consider non-overwhelming forms of coercion. Perhaps some readers will feel that I have built too much into these notions if I have rendered them unavailable to sleep-talkers or to overwhelmingly coerced speakers. However even if, as I deny, standard sleep-talk or forced utterances can be genuine assertions, there is certainly no interesting sense in which they can be lies: someone who accuses me of lying is fully refuted if I can show that I was asleep at the time the speech occurred.

There are indirect ways of holding people responsible for what occurs in sleep, or following overwhelming coercion: we might hold someone responsible for having a will which was too easily overwhelmed, or for allowing herself to be captured, or hypnotised, or drugged, in the first place. Likewise, we might hold someone responsible for having dozed off in public when she knows she is prone to compromising sleep-talk (imagine a politician on the campaign trail). Nevertheless, this doesn't seem like the kind of direct responsibility for speaking falsely which is a precondition of lying: again, if you accuse me of lying, and I can show you that I was asleep when I was speaking, I need no further excuse or justification in this respect, even if I need to apologise for allowing myself to fall asleep in the first place.

Because it is so obvious in these cases that the speaker is not lying, they are not especially interesting for the purposes of this paper. So I now set aside coercion which overwhelms the target's will, to focus instead on the second kind of coercion. This is coercion which changes the relative appeal of the target's options whilst leaving her able to choose between them, whether resentfully or resignedly. Such coercion underpins people's everyday false speech in totalitarian regimes, and it seems to be what Dummett has in mind when writing of capitulations. Moreover, such coerced speech is philosophically interesting for the ways in which it allows us to draw apart different aspects of speakers' responsibility for what is said. The very fact that such coercion is not completely overwhelming indicates that there is some sense in which the speaker is responsible for her choice to speak in that way, but it is also very natural to think that coercion diminishes responsibility in some respect, and/or changes our moral evaluation of the speech in question.

Section 3. Coercion and Normative Practices

Coercion can sometimes excuse behaviour, or justify actions which would otherwise be unjustified. If I trample your flower bed because I am physically pushed onto it, the push means that I had no other option – overwhelming coercion provides me with an excuse. But if I trample the flowers because I am ordered at gunpoint to do so, then the threat is standardly not thought of as an excuse for my behaviour, but rather as a justification for my choosing to act in that way. Trampling your flowers was the right thing to do, given the circumstances, i.e. the potentially fatal consequences of not trampling.

Such coercion is not, however, a 'get out of jail free' card. Whether or not it is justifiable to conform to a coercive threat, for example, depends in part on factors such as how severe the

threat is, and what action is being coerced. I am justified in trampling on your flowers if I am threatened with death otherwise. However it is less clear whether I am justified in trampling on your flowers if my aggressor is merely threatening to trash my garden if I don't trash yours. Even a very serious threat, for example a threat to my own life, might not justify me in committing some horrendous act against others. There is no straightforward moral calculus here, and the justifying power of coercion needs to be assessed case by case.

In addition to these issues around excuse and justification, distinctive issues arise where non-overwhelming coercion is brought to bear upon normative practices. The normative practice of assertion is of course at the heart of the present paper. But I will approach it by first exploring two other normative practices: gift-giving and promise-making. I will argue, as others have done, that such practices can be undermined by non-overwhelming coercion. My main goal is merely to establish that coercion does sometimes have this undermining power, and I do not offer any detailed account of how, why and when this occurs. Nevertheless, I am guided by the accounts of Owens (2012: ch.10) and Pallikkathayil (2011); I will return to each of these authors in course of my discussion.

Imagine that I hand over some of my money to someone who will use it to harm themselves or others. You object to what I have done, and I tell you that I handed over the money only because the person in question threatened me into doing so. The threat means that I did not transfer property rights in handing over the money: if I get the opportunity to take back the money, then I can do so without automatically thereby wronging its 'possessor', whereas genuine gifts cannot be unilaterally retrieved without doing at least *prima facie* wrong to the recipient of the gift. (Sometimes retrieving a genuine gift is all things considered the right thing to do, just as sometimes stealing is all things considered the right thing to do, but nevertheless apologies and compensation are appropriate.)

A genuine gift is not subject to coercion. Thus, if you were objecting to my making a gift of this money, then by citing the threat I show you that I did not make a gift. Relatedly, if you object to my taking back the money when I get the opportunity, then by citing the initial threat I show you that I am not stealing but recovering what was in effect taken from me. (What if I am coerced into handing money to an innocent third party? I return to this issue below.)

We see a similar structure when considering the act of promise-making under duress. Under normal circumstances, an uncoerced promise generates new obligations: the promiser is now obliged to keep the promise. This is not to say that breaking a promise can never be justified: sometimes exigent circumstances can make this the right thing to do. Nevertheless, a broken promise wrongs the promisee. In such exigent circumstances, as David Owens writes, ‘...guilt, remorse, apology, requests for forgiveness are appropriate responses on the part of the promisor even though what they have done [i.e. broken the promise] was justified.’ (2012: 233).

But it is very widely accepted that a (seeming) promise extracted under coercion typically fails to generate such promissory obligations. There are many different views about why coercion has this invalidating power, and indeed about whether coercion always invalidates promise-making, or only typically does so. For example, Owens argues that coercion invalidates promising only if the promisor was wronged by the coercion; he claims that people are not inevitably wronged by coercion, as when parents permissibly coerce their children into better behaviour. John Deigh (2002) argues that during war coerced promises may be binding: under the right circumstances, a coerced surrender may involve a valid promise not to resume fighting. And Seana Shiffrin (2014) distinguishes ‘scripted’ from ‘initiated’ coerced promises. If the target is directly coerced into making a specific, scripted ‘promise’, this does not bind her. But if a coerced target initiates a promise, e.g. as a negotiating move, this can be binding. Both Deigh and Shiffrin stress the importance for all parties of keeping channels of communication and commitment open even under hostile circumstances.

(A note on terminology: some authors write as if valid and invalid promises were both promises. In contrast, I will use scare quotes to indicate invalid ‘promises’, and likewise for coerced ‘gifts’. Nothing of substance is intended to hang on this decision, which will recur below in my discussion of coerced ‘assertions’ and ‘lies’.)

For my purposes it is enough to note the consensus over the point that coercion can very often invalidate a ‘promise’, despite the lack of consensus about the limits and explanation of this phenomenon. If you object to my making a certain ‘promise’, then, in central cases, by citing a threat I show you that I did not genuinely promise. And if you object to my not keeping my ‘promise’, then again in central cases I can defend myself by reference to the

threat, which shows that I am not bound by promissory obligations. (What if I am coerced into making a ‘promise’ to an innocent third party? I return to this issue below.)

Thus coercion can nullify standard normative moves regarding both gift-giving and promise-making. The same is true, notoriously, with coerced ‘consent’. Japa Pallikkathayil (2011) frames this in terms of ‘impaired normative authority’: victims of coercion suffer a reduced ability to make discretionary changes in permissions and obligations. She argues that this captures both the way in which coercion wrongs its victim, and the way in which coercion mitigates the victim’s responsibility for her actions, but without implausibly picturing coerced actions as unchosen, or as nonactions. David Owens (2012: 241-2) argues that the validity of a ‘promise’ depends on whether the speaker’s being able to generate promissory obligations serves her specifically normative interests on that occasion: ‘promises’ extracted under wrongful coercion do not satisfy this condition. Whilst the differences between these accounts (and indeed others) are significant for many purposes, I will not pursue them here: my main point is just to establish that coercion sometimes nullifies ordinary normative moves.

Section 4. But was it Justified?

Even where coercion clearly blocks a normative move in this way, questions of justification may remain. For example, coercion which invalidates a ‘gift’ may not suffice to justify the corresponding act of transfer, given the consequences of handing over the money, and the consequences of refusing to do so.

Imagine that my neighbour threatens to daub graffiti on my front door if I do not provide money to pay for a hit on her ex-husband. I hand over the money, but the threat means that this is not a gift, and so I am not liable to any criticism of my gift-giving practices.

Nevertheless, I am not justified in handing over the money: I have valued my paintwork over a man’s life. In some ways, what I have done is like leaving a pile of cash just inside an open window, knowing that my unscrupulous neighbour is desperate to pay for the hitman. My neighbour does not acquire property rights if she reaches in to steal the money; nevertheless I am criticisable for facilitating my neighbour’s acquisition of the cash, given what she will do with it.

Imagine instead that my neighbour credibly threatens to injure my children if I do not provide enough money to buy hideous ornaments for her front garden, thus lowering the tone of the whole street; I hastily hand over the money. Again, this isn't gift-giving. But in this case, the seriousness of the threat and the relative triviality of the tone-lowering justify my action. I have acted reasonably in valuing my children's safety over neighbourhood aesthetics.

So coercion can prevent a transfer of money or goods from constituting a genuine gift. But it does not always justify the transfer itself, i.e. the act which would have constituted gift-giving had it been uncoerced. Whether the transfer itself is justified depends on factors such as the severity and credibility of the threat, the possible consequences of transferring or not transferring, and the rights and interests of various parties.

The same applies to coerced 'promises'. Suppose my colleague threatens to bad-mouth me to the Dean if I do not promise to mark all the exams for his class, and so I say 'I promise to mark those exams'. My partner is cross when he hears about this, knowing that if I mark the exams our weekend will be ruined. I tell my partner that my 'promise' was extracted via a coercive threat. Now, by most accounts, my partner should appreciate that I have not made a valid promise to my colleague, and that I would not wrong the colleague if I failed to mark the exams. But even accepting that I did not make a valid promise, there is a further question as to whether I was justified in saying 'I promise to mark those exams'.

In saying 'I promise' I may have generated other, non-promissory reasons for me to mark the exams and thus ruin our weekend. The colleague will now not do his own marking, so the students will fail to graduate if I don't step in; others who heard the 'promise' but not the threat will now wrongly but damagingly regard me as a promise-breaker if I fail to do the marking. By saying 'I promise', although I have not promised, I have helped bring about a situation in which I have most reason to do the marking and thus ruin our weekend; my partner may argue that I should have chosen the bad-mouthing rather than the ruined weekend. In addition, perhaps there were reasons not to say 'I promise' which do not turn on the badness of my doing the marking: perhaps I should have stood up to my colleague's bullying ways, either for my own sake or to protect others.

So a coerced 'gift' does not directly result in a transfer of property rights with their concomitant permissions and obligations, and a coerced 'promise' does not generate the standard obligations to act as 'promised'. But, as Pallikkathayil writes,

[where a mugger threatens a victim into 'promising'] the victim's utterance is still morally assessable even though she makes it in order to save her life. If she has an obligation to someone other than the mugger not to engage in deceptive acts, she may have an obligation to follow through with her apparent promise in order to avoid wronging these others. The impaired normative authority account is only committed to claiming that the promise itself is not binding. (2011: 17)

In the cases I have considered so far, coercion restricts the victim's normative authority, as Pallikkathayil would put it, with respect to the victim herself and/or the coercer. That is, the neighbour threatens me into transferring money to *her*, rather than to some innocent third party, and the colleague threatens me into saying 'I promise to mark the exams' to *him*, rather than to the students. But it should be uncontroversial that third party interests may be relevant to whether the coerced person was justified in transferring the money, or in saying 'I promise', even where coercion means that there was no genuine gift or promise. My neighbour's ex-husband's interests are relevant to whether I should have handed over the money that the neighbour will spend on a hitman. And my partner's interests are relevant to whether I should have said 'I promise to mark those exams'. Moreover, as Pallikkathayil suggests, once the initial transfer or utterance has taken place, third party interests may be relevant to whether the coerced person should act as if a genuine gift or promise has been made, even though in fact this is not the case.

Some might be tempted to think that, in addition, under certain circumstances coerced 'gifts' to innocent third parties can constitute genuine gifts and coerced 'promises' to innocent third parties can constitute valid promises. Fully investigating this issue would require me to adopt some more specific view of how coercion nullifies normative moves. But *prima facie* it would be dangerous for us to allow such coerced moves to have the normative status of uncoerced moves; in particular, it would seem wrong for others to enforce the obligations and permissions supposedly created by coercion in this way. Of course, one feels a pull of sympathy to the innocent third parties who are unaware that the 'gift' or 'promise' they have received is not the genuine article. But we can account for this, and for the reasons created

by such innocent third-party expectations, without regarding such reasons as corresponding to property rights or promissory obligations. (Chwang 2011: 162-3).

Section 5. Assertion, Lies, and Normative Authority

How might the idea of impaired normative authority help us understand coerced ‘assertions’? Like promising, gift-giving, and consenting, ordinary assertion is a normative business. This is clear for views which associate assertion with commitment either to defend what is said or else retract it in the light of criticism (Brandom 1983). On such a picture, to make an assertion is to undertake a commitment, imposing new obligations upon oneself, and creating new permissions for the audience. This kind of view is at issue between Dummett, Kenyon, and Sorensen.

Moreover those who work primarily with the notion of telling, rather than assertion, also emphasise normative relationships between speaker and audience (Hornsby 1994, Fricker 2006). For example, Hinchman (2005) argues that telling involves an invitation to trust, i.e. the creation of a permission to trust, and, plausibly, an obligation upon oneself to be trustworthy in the relevant respect. Similarly, Moran (2005) understands testimony in terms of the offering of assurance, a taking of responsibility for what others may believe.

Intriguingly, Moran makes a connection between assertion and freedom of choice:

Further, a specific assumption of responsibilities is essentially an expression of a person’s freedom, something that only makes sense as consciously assumed. It is for this reason that words spoken during sleep or under hypnosis do not have the value of testimony, because they do not count as assertions, whatever expressive psychological value they might still retain as evidence. Like a promise or an apology, something only counts as a person’s assertion when consciously presented as such by him.
(2005: 18)

But one can recognise normative aspects of assertion without adopting either an assurance picture or a Brandom-style ‘commitment account’. It is very widely thought that asserting involves taking responsibility, making oneself accountable, or undertaking something: this is typically taken as a starting point, before different authors diverge in their analyses of quite how this works. For Peirce, ‘to assert a proposition is to make oneself responsible for its truth’ (1932: 384). For Searle an assertion of p ‘counts as an undertaking to the effect that p

represents an actual state of affairs' (1969: 66). For Alston, an assertion that *p* involves a speaker's taking responsibility for its being the case that *p* (2000: 120); more generally, Alston understands a whole range of 'illocutionary acts' in terms of the speaker's taking responsibility for the satisfaction of relevant conditions. Even for Williamson, '[t]o make an assertion is to confer a responsibility (on oneself) for the truth of its content' (2000: 268), a responsibility discharged by knowing the truth of that content, and thus satisfying what Williamson identifies as the epistemic norm governing assertion.

This standard connection between asserting and taking responsibility may partially explain what is so disorienting about some of Donald Trump's practices of speech and tweeting (I write in the run-up to the 2017 inauguration). He speaks in contexts which are appropriate for assertion (e.g. not in theatrical contexts), with the mannerisms and tone of someone who is asserting, yet sometimes seems oblivious to any sense that such speech incurs commitments. In line with my arguments so far, one might wonder whether this is really a practice of assertion at all. To begin to explore this, we would need to consider questions about the relative rôles of speaker intentions and audience expectations in determining the significance of a speech practice (Kukla 2014). It is important to acknowledge that speech can be morally evaluable in connection with its truth value even if it does not constitute assertion: even in more ordinary situations, acts of hinting, speculating, questioning, and so on are all morally evaluable.

I have invoked a tangle of multifarious norms; different accounts of assertion, telling, and even testimony will articulate the normative aspects in different ways. But however this is spelt out, we should expect to find respects in which, under ordinary non-coercive circumstances, a person who tells someone that *p* exercises her normative authority, in Pallikkathayil's sense, in a way which can be curtailed by coercion. An uncoerced teller creates obligations upon herself – perhaps to speak knowledgeably, or at least to avoid deliberate falsehood – and creates permissions for her audience, perhaps to believe that *p* on this basis, or to hold the speaker responsible in some sense for the truth of *p*.

We saw earlier that coercion can restrict the normative authority of its target. Just as coerced 'gifts' do not transfer property rights, and coerced 'promises' do not generate promissory obligations, coerced 'assertions' do not generate the standard normative shifts associated with uncoerced assertions. So for example, one cannot be coerced into taking responsibility for

the truth of p , in the relevant sense of making oneself liable to criticism if p is false at the moment of speaking.

It is true that people can be coerced into doing things which result in their becoming obliged to make it the case that p , for some p . For example, when I am coerced into saying ‘I promise to mark the exams’, making the students dependent upon me for their graduation, then perhaps I become obliged to mark the exams, even though I did not genuinely promise to do so. Then in a sense I have been coerced into taking responsibility for the truth of a proposition, the proposition that the exams will be marked. But this responsibility-for-making-true is not the type of responsibility undertaken in (uncoerced) assertion according to Peirce, Searle, Alston or Williamson. Rather, in asserting, one makes oneself somehow answerable for p ’s being true as you assert it. It is this making oneself answerable that is undermined by coercion, just as the normative shifts associated with ordinary promising and gifting are undermined by coercion.

Are coerced ‘assertions’ really assertions at all? My claim is that they do not have the normative significance of uncoerced assertions, and in line with my earlier practice I will use scare-quotes: they are mere ‘assertions’. An alternative practice would be to refer to such coerced utterances as invalid or nullified assertions, but nothing of substance hangs on this. Likewise, *a fortiori*, coerced ‘assertions’ are not genuine lies. We could call them invalid or nullified lies, but again, nothing of substance hangs on this.

Section 6: Justification

Coerced ‘lies’ are not genuine lies. But the coerced speaker is not automatically justified in making insincere utterances. As with ‘gifts’ and ‘promises’, we can evaluate a speaker’s decision to make an ‘assertion’ under coercion: non-overwhelming coercion does not automatically justify every action. I will consider a few cases to illustrate how such evaluations might proceed. But we should bear in mind the wide variety of possible situations in which speech may be coerced. For example, it may make a difference whether the audience and the coercer are one and the same, what the audience knows about whether the speech is coerced, and whether the speaker is coerced to say that p , or else to say sincerely whether p . Other relevant factors include the degree of coercion applied, and the consequences of complying or not complying. As I noted earlier in this paper, there are many

cross-cutting ways of taxonomising cases of coerced speech, and not every distinction is a sharply determinate one.

Let's begin with capitulation-style coercion to say that p , to an audience which perpetrates and thus knows about the coercion: the case of Dummett's persecuted Christian, for example. Given the coercion, such an utterance is not an assertion, and *a fortiori* not a lie.

Nevertheless, we can weigh up the reasons to make this utterance against the reasons not to make it, and try to make a judgement about whether the coerced's decision to say that p was justified. Although of course this calculation will vary enormously with the details of the case, it is striking that the truth of p , and whether the speaker believes that p , do not look like especially weighty factors in the balance, given that the audience knows of the coercion. It is not obvious that there is always some reason to prefer sincere utterances in such cases: much depends upon whether sincerity has better or worse psychological or practical consequences for the speaker. Some people might strongly resist insincerity – perhaps especially in religious contexts – but others might feel that it is more dignified not to express their genuine beliefs in such circumstances.

What about everyday life in the totalitarian state, in which the audience knows of the coercion, so there is no attempt at mutual deception, but the audience is not responsible for the coercion? Again, I argue that the coercion means that the relevant utterances are not genuine assertions, but nevertheless there are questions to be asked about whether speaking this way is justifiable or not. For example, we might weigh up the severity of the threat, the likelihood that any rebellion would spread and succeed, whether there are other priorities for resistance, and so on. The fact that the audience for such assertions is not misled into thinking them sincere, and moreover does not suffer the insult or disrespect inflicted by a genuine lie, presumably weighs in favour of such utterances being justified. But it is an empirical question whether even patently insincere utterances can cumulatively give legitimacy to what is said, contributing to an ethically-significant process of normalisation.

What difference does it make if such known-to-the-audience coercion is directed not at getting the person to say that p , but at getting them to say whichever of p or *not- p* they believe? (Again, I will take it that the coercion prevents the utterance from being a genuine assertion.) There is a sense in which the speaker has more options available to her in such situations. When a speaker is coerced to say that p , then under most circumstances it is

obvious to the coercer whether or not the speaker does what the coercer requires. So the speaker has only two main choices: to accede and say that p , or to defy the coercion openly. When a speaker is coerced to say whichever of p or $not-p$ they believe, it may not be obvious to the coercer whether or not the coercion has succeeded. The speaker can choose defiant silence, she can accede and say whichever of p or $not-p$ she believes, or she can say whichever of p or $not-p$ she does not believe, but try to appear sincere nevertheless. Depending on the chances of being found out, however, the potential cost of this third option may be even greater than the costs incurred by remaining silent.

What about coercion to say that p , or to reveal whether p , to an audience which is not aware of the coercion? I have already suggested that coercion undermines its target's ability to undertake obligations even to innocent third parties, so I maintain that a coerced speaker is not making an assertion even under such circumstances. But the moral innocence and epistemic ignorance of a third-party audience will weigh in the balance in considering whether such utterances are justified; as ever, of course, this will depend upon the details of the case in question. (Michaelson (2016) discusses whether a coerced speaker who tries to get an innocent audience to accept that p can count as intentionally deceiving the audience.)

In military or security contexts, a 'duress code' or 'duress signal' is sometimes established, so that, for example, hostages are able to signal to a friendly audience that they are speaking under duress, without alerting or provoking their captors. The signal can involve a fixed verbal phrase, or a nonverbal code such as crossed fingers. (The captured crewmen of the USS Pueblo in North Korea improvised various such signals: Bucher and Rascovich (1970).) One important feature of these signals is that they lower the justificatory threshold for speaking falsely under coercion, by lowering the chance that a morally innocent third-party audience will be deceived.

Section 7. Conclusions

I have argued that once we recognise that assertion or telling involves normative powers, and that non-overwhelming coercion can impair our normative authority, then we should recognise that coercion can prevent an utterance from being an assertion, and *a fortiori* from being a lie, even though it has the various surface features which normally indicate assertion. In the absence of a sharp distinction between noncoercive and coercive pressure, this doesn't

provide us with a sharp distinction between lies and ‘lies’, but there will be plenty of cases which are uncontroversially coercive.

As with invalid ‘promises’ and coerced ‘gifts’, however, coercion does not automatically settle whether the coerced person is justified in behaving in a certain way: we can still assess intentionally false speech along a moral dimension, even once we recognise that it does not constitute lying. We still have some responsibility for the ways in which we behave under non-overwhelming coercion. A full exploration of the many ways in which speech can be coerced, and the many factors relevant to the justification of coerced speech, would amount to a lengthy field-guide, which I have not attempted here.

Section 8. Afterword - Tricked into Lying?

Duress is often discussed alongside deception, especially in legal contexts. Just as consents and contracts extracted under duress are regarded as invalid, consents and contracts obtained through deception likewise do not have the significance they would otherwise have. As with duress, deception can take many forms, and many degrees of severity; mild deception or failure to disclose does not always invalidate a consent or contract. Nevertheless, deception can sometimes have the power to invalidate. What, then, should we say about seeming-assertions which are obtained via deception: is it possible to trick someone into making an assertion, indeed trick someone into telling a lie?

When we think about these issues with regard to gifting, promising, or indeed consenting, a typical scenario is as follows. Someone is deceived into transferring some goods, or saying ‘I promise’, or ‘I consent’. It then becomes apparent to this person that she has been deceived, and the question arises whether she is permitted to recover the cash, break the ‘promise’, or withdraw her ‘consent’. (Correlatively, is the deceiver now entitled to behave as if she has property rights, or the permissions normally associated with having received a promise or a consent?) That is, we typically imagine a time-lag between the moment at which the person is deceived into taking the initial action, and the moment at which, now enlightened, she must decide whether to assign normative significance to her earlier action.

Deceptively obtained assertions do not typically involve a time lag in this way. For the most part, the obligations one undertakes by asserting are discharged (or not) in the very moment of speaking. An obligation to speak sincerely with regard to whether *p* is met or violated

immediately it is incurred by asserting as to whether p . Likewise for the putative obligation to speak knowledgeably. If Brandom and others are right in thinking that assertion involves commitments which bind us at later times – commitments to defend or retract, for example – then this does introduce the possibility of time lag. But such views do not entail that assertion does not also make normative demands upon the speaker at the moment of speech.

Someone who is deceived into making an assertion-like utterance believes as she speaks that she is genuinely asserting, even if in fact the deception makes it the case that she does not genuinely assert. Suppose that someone in these circumstances speaks falsely, with the intention to deceive. Then, insofar as we are judging the speaker, rather than her action, the speaker seems as culpable for this speech as she would have been were she genuinely lying, as indeed she takes herself to be. That said, so long as we do not take an exceptionlessly strict line against lying, the nature of the deception will be relevant to our evaluation here: someone who is deceived into ‘lying’ for what she thinks is the greater good is more admirable than someone who is deceived into ‘lying’ for what she thinks will be her personal gain.

Just as we saw with coercion, this very brief discussion of deceptively-obtained ‘lies’ indicates the importance of attention to the varying details of different cases. Nevertheless, as with coercion, this again illustrates the usefulness of distinguishing between the power of deception to undermine the usual normative structure of assertion, and the more limited power of deception to excuse or justify certain types of speech.

Acknowledgements

This paper was supported by a Leverhulme Major Research Fellowship, which I gratefully acknowledge. The material was presented at the Irish Philosophical Club in Drogheda, at the University of Liverpool, at the 2015 Joint Session in Warwick, at the University of Glasgow Philosophy Society, and at a St Andrews reflectorium; my thanks go to all those audiences for lively discussions. Thanks in particular to Garrett Cullity, Tom Dougherty, Tim Kenyon, Ishani Maitra, Rachel McKinney, Eliot Michaelson, Japa Pallikkathayil, Andreas Stokke, and Bill Wringe, for very useful conversations and/or written comments. Thanks also to the referees for this volume, both for their philosophical input and for the reference to Endō’s novel *Silence*.

References

- Alston, W. P. (2000), *Illocutionary Acts and Sentence Meaning* (Cornell University Press).
- Anderson, S. (2015), 'Coercion', *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2015/entries/coercion/>.
- Brandom, R. (1983), 'Asserting', *Noûs* 17/4: 637–650.
- Brudney, V. (2000), 'Compelled Speech', in L.W. Levy, K.L. Karst, and A. Winkler (eds.) *Encyclopedia of the American Constitution* (Macmillan Reference USA). *****no page references since I consulted it online: retrieved online January 11th 2017 <http://www.encyclopedia.com/politics/encyclopedias-almanacs-transcripts-and-maps/compelled-speech>*****
- Bucher, L.M. and Rascovich, M. (1970), *Bucher: My Story* (Doubleday).
- Carson, T. (2006), 'The Definition of Lying', *Noûs* 40/2: 284-306.
- Chen-Wishart, M. (2015), *Contract Law 5th Edition* (Oxford University Press).
- Chwang, E. (2011), 'On Coerced Promises', in H. Sheinman (ed.) *Promises and Agreements* (Oxford University Press) 156-82.
- Deigh, J. (2002), 'Promises Under Fire', *Ethics* 112/3: 483-506.
- Dummett, M. (1981), *Frege: Philosophy of Language* (2nd edition) (Duckworth, and Harvard University Press).
- Fricker, E. (2006), 'Second-Hand Knowledge', *Philosophy and Phenomenological Research*, 73/3: 592-618.
- Gilbert, M. (1993), 'Agreements, Coercion, and Obligation', *Ethics* 103/4:679–706.
- Grice, H.P. (1957), 'Meaning', *Philosophical Review*, 66/3: 377-88.
- Hinchman, E. (2005), 'Telling as Inviting to Trust', *Philosophy and Phenomenological Research*, 70/3: 562-87.
- Hornsby, J. (1994), 'Illocution and its Significance', in S.L. Tsohatzidis (ed.), *Foundations of Speech-Act Theory: Philosophical and Linguistic Perspectives* (Routledge): 187-207.
- Kenyon, T. (2003), 'Cynical Assertion: Convention, Pragmatics, and Saying "Uncle"', *American Philosophical Quarterly* 40/3: 241–247.
- Kenyon, T. (2010), 'Assertion and Capitulation', *Pacific Philosophical Quarterly* 91/3: 352-368.
- Kukla, R. (2014), 'Performative Force, Convention, and Discursive Injustice', *Hypatia* 29/2: 440-54.

- Leland, P.R. (2015), 'Rational Responsibility and the Assertoric Character of Bald-Faced Lies', *Analysis* 75/4: 550-54.
- Maitra, Ishani (this volume)
- Maitra, I. and McGowan, M.K. (eds.) (2012), *Speech and Harm: Controversies over Free Speech* (Oxford University Press).
- McKinney, R. (2016), 'Extracted Speech', *Social Theory and Practice*, 42/2: 258-84.
- Michaelson, E. (2016), 'The Lying Test', *Mind and Language* 31/4: 470-499.
- Moran, R. (2005), 'Getting Told and Being Believed', *Philosophers' Imprint* 5/5: 1-29.
- Owens, D. (2012), *Shaping the Normative Landscape* (Oxford University Press).
- Pallikkathayil, J. (2011), 'The Possibility of Choice: Three Accounts of the Problem with Coercion', *Philosophers' Imprint* 11/16: 1-20.
- Peirce, C.S. (1932), *Collected Papers of Charles Sanders Peirce* vol 2, edited by C. Hartshorne and P. Weiss, (Harvard University Press).
- Perillo, J.T. and Kassin, S.M. (2011), 'Inside Interrogation: The Lie, the Bluff, and False Confessions', *Law and Human Behavior* 35/4: 327-37.
- Saul, J. M. (2012), *Lying, Misleading and What is Said* (Oxford University Press).
- Searle, J. (1969), *Speech Acts: An Essay in the Philosophy of Language* (Cambridge University Press).
- Shiffrin, S.V. (2014), *Speech Matters: On Lying, Morality and the Law* (Princeton University Press).
- Shue, H. (1978), 'Torture', *Philosophy and Public Affairs* 7/2: 124-143.
- Sorensen, R. (2007), 'Bald-Faced Lies! Lying Without the Intent to Deceive', *Pacific Philosophical Quarterly* 88/2: 251-264.
- Williamson, T. (2000), *Knowledge and its Limits* (Oxford University Press).