

INNER SPEECH: A PHILOSOPHICAL ANALYSIS

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF  
PHILOSOPHY AT THE AUSTRALIAN NATIONAL UNIVERSITY

DANIEL JOHN GREGORY

MAY 2017

© Copyright by Daniel John Gregory, May 2017

All rights reserved

This thesis is entirely my own original work.

Word count, excluding footnotes and appendix: 50,403.

Word count, including footnotes and appendix: 60,056.

*Daniel Gregory*

.....

Daniel Gregory

*For my parents*

## *Acknowledgements*

Every PhD student has some help along the way. I have had more than most.

First, I thank my primary supervisor, Daniel Stoljar. Daniel is devoted to his students; he prioritises their work above his own. He has taught me the importance of detail in philosophy. With his help, I have not written the best dissertation that I thought I could; I have written a better dissertation than I thought I could.

Next, I thank my associate supervisors, David Chalmers, Frank Jackson and Kim Sterelny.

Dave teaches, in large part, by asking difficult questions, so many of which open up new possibilities. He also knows just how much to challenge students; he does not dismantle their work.

Frank can take a paper and provide a single, brief insight which illuminates everything. He also models the way an intellectual should conduct him- or herself. I admire his humility as much as his mind.

I did less work with Kim than I initially envisaged, as the project ultimately did not have so strong an empirical aspect as I once thought it might, and for which reason I sought his expertise. Still, he taught me how to approach a new and overwhelming literature outside my own area, and gave me confidence in doing so, even if that aspect of my research was gradually deemphasised.

I doubt that there is a graduate student in philosophy anywhere in the world who could confidently claim to work with a more eminent group of supervisors. I am aware of the privilege I have had.

Others at the Australian National University who have helped in one way or another include Alma Barner, Zoe Drayson, Anya Farennikova, Brian Garrett, Liz Irvine, Robert Kirby, Colin Klein, Victoria McGeer, David Nash, Luke Roelofs, Alan Rumsey, Jonathan Simon and Inga Vermuelen.

Between February and May 2014, I was a visiting student at the University of Oxford. I am grateful to Martin Davies for facilitating my visit there and for his detailed feedback on my work. I am also grateful to Philipp Koralus for discussing my work.

While based at Oxford, I made some shorter visits to the Hearing the Voice project at Durham University and the Centre for the Study of Perceptual Experience at the University of Glasgow. For their hospitality, discussions and feedback, I am grateful to Ben Alderson-Day, Charles

Fernyhough, Matthew Ratcliffe, David Smailes and Sam Wilkinson at Durham, and to David Bain, Mike Brady, Jennifer Corns and Fiona Macpherson at Glasgow.

Between July 2015 and February 2016, I was a visiting student in the School of Mind and Brain at Humboldt University of Berlin. I am grateful to Michael Pauen for facilitating my visit and for discussing my work, and to Lena Kästner, Richard Moore and Anna Strasser for discussions and comments on my work. I am also grateful to a larger group there for making me so welcome: Frederik Bartels, Maximilien Chaumon, Sanneke de Haan, Joerg Fingerhut, Luca Iemi, Maurício Martins and Lara Pourabdollahim, as well as Lena and Richard.

Others who have helped include Joanna Atkinson, Tyler Burge and John Maier. I am especially grateful to Hugh Mellor for discussion and comments on a draft paper.

A significant amount of the work in this dissertation has been published and it has been improved on account of comments from referees during the review process. Some of the material in Chapter One and most of the Appendix has been published as ‘Inner Speech, Imagined Speech, and Auditory Verbal Hallucinations’ (2016) *Review of Philosophy and Psychology*, vol. 7, no. 3, pp. 653-673; most of the material in Chapter Two has been published as ‘Is Inner Speech Dialogic?’ (2017) *Journal of Consciousness Studies*, vol. 24, no. 1-2, pp. 111-137. I am grateful to the referees who reviewed these papers, including Charles Fernyhough, who identified himself as a referee of the second paper.

Parts of this dissertation or related material were presented to audiences at an ANU graduate conference at Kioloa in November 2012; at ANU Philsoc seminars in March 2013 and August 2014; at the Serious Metaphysics Group at the University of Cambridge in April 2014; at Michael Pauen’s colloquium at Humboldt in October 2015; and at the ANU Philosophy of Mind Work-in-Progress Group in June and November 2016. I am grateful for the feedback I received at each of these venues.

I am grateful to Dominie Dessaix, Jennifer Gregory and Tim Williamson for proofreading parts of the dissertation.

I may well have forgotten others who have helped at one point or another during my candidature. I regret any omissions.

Since commencing my candidature, I have been the recipient of an Australian Postgraduate Award funded by the Commonwealth Government; a University Winter Course Scholarship from the German Academic Exchange Service; an ANU Vice-Chancellor’s Travel Grant; an ANU-IARU (International Alliance of Research Universities) Travel Grant; a Research Grant for Doctoral Candidates and Young Academics and Scientists from the German Academic

Exchange Service; and funds from the Department of Philosophy at ANU supporting travel and attendance at conferences and a summer school.

I now move on to some personal notes.

The following people – fellow PhD students, ANU faculty, friends and family – have all supported me in one way or another during the course of my candidature: Alma Barner, Rachael Brown, Justin Bruner, Clare Due, Jonathan Farrell, Bronwyn Finnigan, Clare Gregory, James Gregory, Pascal Herington, Rachel Kirby, Robert Kirby and Rosa Terlazzo. There are many others.

There have been some very difficult times since just before I came to Canberra to commence my PhD in 2012. There are many people to whom I am deeply indebted for their help in getting through these times, most especially Brian Gray, Anna Gregory, Kate Horning, Helen Irving and Elisabeth Smartt. I will not describe here what they have done for me, but I hope they know that I am grateful with my whole heart.

I want to record my gratitude to my grandparents and to one dear aunt. Both of my grandfathers, Edward James (Jim) Lynch and John (Jack) Gregory, have passed away; my grandmothers, Patricia Lynch and Katherine Gregory, are still with us. All have been constant in their love and encouragement.

Christine Lynch has been devoted to her nieces and nephews. We are so fortunate that she has been so involved in our lives.

Finally, I thank my parents, Jennifer and Peter Gregory, to whom this dissertation is dedicated. They have supported me in every way. I could not have completed the dissertation without them. I am humbled by their love.

## **Abstract**

This dissertation explores the phenomenon of inner speech. It takes the form of an introduction, which introduces the phenomenon; three long, largely independent chapters; a conclusion; and an appendix.

The first chapter deliberates between two possible theories as to the nature of inner speech. One of these theories is that inner speech is a kind of actual speech, just as much as external speech is a kind of actual speech. When we engage in inner speech, we are actually speaking, but we are doing so silently. The other theory holds that inner speech is a kind of imagined speech. When we produce inner speech, we are imagining performing the action of speaking. The chapter argues for the theory that inner speech is a kind of actual speech.

The second chapter argues against a theory which holds that inner speech is dialogic. On this theory, a subject represents different perspectives in inner speech and a dialogue can take place in the same sense in which a dialogue can take place between different individuals in external speech. The chapter borrows some important material from the philosophy of language to show that this position, though it might have some intuitive appeal, is ultimately implausible.

The third chapter is concerned with the question whether inner speech can be a source of knowledge of our own beliefs. It shows that the view that inner speech can be such a source is subject to an adapted version of a problem from the epistemology of testimony: roughly, what justification do we have for believing that we believe what we say in inner speech? It makes use of some material from the recent debate about cognitive phenomenology to develop a version of the view which is not subject to this problem. It then provides some initial discussion of the merits of this view.

The appendix takes up a more practical issue regarding inner speech. There is a theory that auditory verbal hallucinations – i.e. experiences of voice-hearing – take place when someone produces an utterance in inner speech but loses track of the fact that they have produced the utterance. Accordingly, they have an experience as of something being said and, not realising that they are the source of the experience, postulate some external cause, i.e. someone else speaking. The appendix develops an alternative account which has been suggested in the literature, at times drawing upon earlier work in the dissertation.

## Contents

Introduction.....	12
1. The Project.....	12
2. Finding the Target.....	14
3. Some Examples.....	17
4. Vygotsky and the Development of Inner Speech.....	20
Chapter One: Imagined Speech or Actual Speech? .....	23
1. The Problem.....	23
2. The Possible Views.....	24
2.1. The Imagined Speech View .....	25
2.2. The Actual Speech View.....	30
2.3. Quasi-Speech .....	34
3. Inner Speech is Actual Speech.....	35
3.1. The Embedding Argument.....	35
3.1.1. Stating the Argument .....	36
3.1.2. Premises 1 and 2 .....	36
3.1.3. Premise 3.....	37
3.1.4. A Refinement .....	37
3.1.5. Two Lingering Concerns .....	38
3.2. The Parallel Cases Argument.....	40
3.2.1. Two Cases .....	40
3.2.2. Possible Speech.....	41
3.3. The Continuity Argument .....	42
3.3.1. Continuous Phenomena.....	42
3.3.2. Imagination and Performance .....	43
3.4. The Precisification Argument .....	43
3.4.1. Suggestion 1 .....	45
3.4.2. Suggestion 2.....	45
3.4.3. Suggestion 3.....	46
3.4.4. Suggestions 4 and 5.....	47
3.4.5. Suggestion 6.....	49
3.4.6. Suggestion 7.....	50
3.4.7. Completing the Argument.....	51
4. Concrete Tokens .....	52
4.1. The Concern.....	52



4.2.	Two Difficulties .....	53
4.3.	Meeting the Challenge .....	54
4.3.1.	Non-Physical Tokens .....	54
4.3.2.	Intuitive Judgements .....	55
4.3.3.	Linguistic Images .....	56
4.3.4.	Combining the Responses .....	56
5.	Assessing the Case .....	56
5.1.	The Two Views .....	57
5.2.	Reviving Imagined Speech? .....	58
6.	Some Implications .....	59
6.1.	Internalisation .....	59
6.2.	Other Such Phenomena? .....	60
6.3.	Aptness for Moral Judgement .....	61
Chapter Two: Is Inner Speech Dialogic? .....		62
1.	The Problem .....	62
2.	The Dialogue Theory .....	64
2.1.	From Interpersonal to Intrapersonal: Some Context .....	64
2.2.	What is Dialogue? .....	65
2.3.	Abbreviation .....	70
3.	Assessing the Case .....	71
3.1.	Representing Perspectives .....	71
3.2.	Representing Perspectives in Dialogue .....	73
4.	Against the Dialogue Theory .....	76
4.1.	Grice and Cooperation .....	77
4.2.	Cooperation and Inner Speech .....	80
4.3.	Developments in Gricean Theory .....	81
4.3.1.	Martinich and Authenticity .....	81
4.3.2.	Kasher and Rationality .....	82
4.3.3.	Sperber & Wilson and Relevance .....	83
4.4.	An Analogy .....	85
4.5.	Virtual Interlocutors .....	86
5.	Some Objections .....	87
5.1.	Inner Speech and Self-Deception .....	87
5.2.	Self-Knowledge .....	88
5.3.	Dialogue without Rules .....	89
6.	An Alternative View .....	90
7.	Some Implications .....	90

7.1.	Conditionals .....	91
7.2.	Inner Speech and Accessibility Problems .....	93
Chapter Three:	Inner Speech and Self-Knowledge.....	95
1.	Introduction.....	95
2.	Inner Speech Interpretationism .....	96
3.	A Problem .....	100
3.1.	Vulnerability .....	100
3.2.	Vulnerability and Inner Speech.....	101
3.3.	Borrowing a Solution? .....	103
3.3.1.	Ross and Rules of Language.....	104
3.3.1.	Burge and the Acceptance Principle .....	105
4.	A Detour.....	109
4.1.	Cognitive Phenomenology .....	109
4.2.	Speech Acts.....	111
5.	The Phenomenology of Speech.....	114
6.	Supplemented ISI.....	116
6.1.	Introducing Supplemented ISI .....	116
6.2.	Phenomenology and Justification .....	118
6.3.	The Application of Supplemented ISI.....	122
7.	Dialogue Revisited.....	122
8.	Some Evaluation .....	123
8.1.	Some Strengths .....	123
8.1.1.	Reconciling Intuitions .....	123
8.1.2.	Uncertainty.....	124
8.1.3.	Actual Speech and Imagined Speech Again .....	124
8.2.	Some Weaknesses .....	125
8.2.1.	Assumptions.....	125
8.2.2.	Other Sources of Self-Knowledge .....	125
8.2.3.	Notions of Justification.....	126
Conclusion	.....	127
Appendix:	Inner Speech and Auditory Verbal Hallucinations .....	131
1.	Introduction.....	131
2.	The Inner Speech Theory.....	132
2.1.	The Basic Version.....	132
2.2.	Self-Monitoring and Agency .....	134
2.3.	Further Developments.....	135
2.4.	Putting it Together .....	136

3. Hearing Voices.....	137
4. A Problem for the Inner Speech Theory .....	138
4.1. Inner Speech and Imagined Speech Again.....	138
4.2. An Objection.....	140
5. Advantages of the Imagined Speech Theory .....	141
6. Conclusion .....	143
Reference List.....	145

# Introduction

## 1. The Project

Until very recently, philosophy had neglected inner speech. It had paid a very great deal of attention to the nature of the mind, and a very great deal of attention to the nature of language and speech, but very little attention to the conscious occurrence of speech in the mind. This has changed just in the last few years, with a small number of philosophers beginning to investigate the phenomenon.<sup>1</sup> In this dissertation, I am going to address four questions about inner speech.

First: What exactly is inner speech? What are we doing when we produce inner speech? There seem to be two possibilities. One is that inner speech is a kind of actual speech, just as much as external speech is a kind of actual speech. When we engage in inner speech, we are actually speaking, but we are doing so silently. The other possibility is that inner speech is a kind of imagined speech. When we produce inner speech, we are imagining performing the action of speaking. Which of these is the correct characterisation of inner speech?

Secondly: Does inner speech sometimes take the form of a dialogue? There is a theory in psychology which holds that it does. As we will soon see in a little more detail, the most influential thinker about inner speech, Lev Vygotsky, observed that children first encounter speech in social contexts. They are spoken to by their caregivers, and they witness conversations between other people in their environment. He observed that, as young children, they then begin to produce a type of external speech – egocentric speech – in which they seem to make remarks about whatever activities are occupying their attention (Piaget had observed this as well). Vygotsky suggested that, a little later, this practice is internalised and becomes inner speech. More recently, it has been suggested that inner speech (and, indeed, the egocentric speech which precedes it) not only develops from children’s experience of conversational exchanges, but also that it retains the character of these conversational exchanges – i.e. that inner speech itself is dialogic. It is believed that a subject represents different perspectives in inner speech and that a dialogue takes place in the same sense in which a dialogue takes place between different individuals in external speech. The view does

---

<sup>1</sup> Contributions include Machery (2005), Clowes (2007), Martínez-Manrique & Vicente (2010), Vicente & Martínez-Manrique (2011), Jorba & Vicente (2014), Langland-Hassan (2014), Gelfert (2015), Roessler (2016), Knappik (forthcoming). A workshop on inner speech, *Inner Speech: Theories and Models*, involving several philosophers, took place in Granada in 2015. Peter Langland-Hassan & Agustín Vicente are preparing an anthology on inner speech which will include contributions from several philosophers (the working title is *Inner Speech: Nature, Functions, and Pathology*). Langland-Hassan is also preparing a book on inner speech for the Routledge *New Problems of Philosophy* series.

have some introspective appeal. It does sometimes seem apt to describe inner speech as a kind of internal dialogue. But the view also seems very odd in light of one obvious fact about inner speech: there is only one individual who produces utterances in inner speech. Is inner speech dialogic? That is, does inner speech actually involve a kind of internal conversation? Or does it just sometimes seem like it?

Thirdly: Is inner speech a source of knowledge about our own beliefs? Again, the idea that it is such a source has a lot of introspective appeal. Often, it at least seems that we learn what we think by listening to what we say in our minds. But the idea is subject to an adapted version of a problem from the epistemology of testimony. With respect to interpersonal speech, it is natural to ask what justification we have for believing what others say. With respect to inner speech, we can just as well ask what justification we have for believing that we believe what we say. Can the view that inner speech is a source of knowledge of our beliefs be developed in a way that provides an answer to this justificatory question?

Fourthly: Is there a connection between inner speech and auditory verbal hallucinations (AVHs), i.e. experiences of ‘voice-hearing’? An influential theory holds that there is. On this theory, a subject experiences an AVH when they produce an utterance in inner speech but lose track of the fact that they have produced the utterance. When this happens, they have an experience as of something being said and, not being aware that they have produced the utterance which they seem to be hearing, postulate an external cause for it. That is, they suppose that they are hearing someone else speaking. But there are some oddities about this theory. Most obviously, inner speech typically occurs in a subject’s own voice, whereas AVHs typically occur in the voice of someone other than the subject. Can this theory be improved upon?

Each of the first three questions will be treated in a paper-length chapter. The chapters are largely independent, though connections between them will be noted from time to time. A conclusion will show that the positions argued for in the various chapters are consistent, and also note some further connections. The final question is of a somewhat more practical nature than the others and will be addressed in an appendix.

Overall, then, my primary aim in this dissertation is to contribute to the nascent philosophical literature on inner speech. A secondary objective is to show how investigating the phenomenon of inner speech can cast light on various other topics.

The first thing that needs to be done, though, is to make sure that the right phenomenon is in view. In the next two subsections, I will itemise some characteristics which can, I expect, be attributed to inner speech without too much controversy, and then offer a wide range of examples of the phenomenon. I will then provide an overview of Vygotsky’s highly influential

theory on the development of inner speech. It will be worthwhile to summarise Vygotsky's theory at this early stage, as I will return to it at various points in the dissertation.

## **2. Finding the Target**

What is inner speech? As will become clear, especially in the first and second chapters of this dissertation, there is some room for disagreement here. Still, there is a small number of features which it should be possible to attribute to the phenomenon with reasonable confidence.

First, an instance of inner speech takes place in the mind. It is a psychological occurrence, not something that takes place in the external world. We have a sound-like experience of our own inner speech, but we do not hear anyone else's inner speech. Secondly, an instance of inner speech is a linguistic occurrence. It involves semantic tokens arranged according to some syntax, where these are the words and grammatical rules of a human language, such as English or German.<sup>2</sup> (Or, at the very least, producing inner speech involves imagining producing semantic tokens arranged according to a syntax; this point will be settled in Chapter One). Probably, anyone who turned their mind to the nature of inner speech would take these first two points to be so obvious as to go without saying. Still, it would be an omission not to record them.

Thirdly, inner speech is episodic. A distinction is often drawn between standing mental states and episodic mental states. A belief is a paradigm example of a standing mental state. If you believe that a square is a polygon with four sides of equal length and four equal angles, then you believe that a square is a polygon with four sides of equal length and four equal angles whether or not you are thinking about it at any particular time. You believe it when you are thinking about lunch and you believe it when you are asleep. This is different from, say, visual imagery, which is episodic. If you visualise an orange square at a particular time, you will most likely not still be visualising an orange square later when you are thinking about lunch or when you are asleep. The image is only temporary. An utterance in inner speech is episodic; it is an event which has a beginning and an end, just as an utterance in external speech does. Of course, it may sometimes be a little difficult to say when one episode of inner speech stops and another starts; judgements about this might be somewhat arbitrary.

Next, an instance of inner speech involves the generation of mental imagery. Typically, inner speech involves auditory imagery. There is a sound-like experience, an experience as of

---

<sup>2</sup> Or loosely so. Obviously, we do not carefully adhere to the rules of grammar in inner speech. There will be more on this momentarily.

hearing speech. There is an exception: some deaf people who know a sign language engage in inner signing. However, we know very little about the imagery that inner signing involves.<sup>3</sup> For the purposes of this dissertation, I am going to set inner signing aside. A study of that phenomenon would require specialist expertise and I am not going to assume that what I say about inner speech would automatically apply to inner signing.

The fact that inner speech involves auditory imagery is relevant to the next point: inner speech can be used in the service of working memory, i.e. very short term recall of some kinds of perceptual information. On the influential account described in Baddeley (1992), one component of working memory is a ‘phonological loop’. This phonological loop comprises a ‘phonological store’ (p. 255) which can briefly retain auditory imagery and an ‘articulatory control process’ (p. 255) which allows one to produce silent utterances. When one produces a silent utterance – i.e. an utterance in inner speech<sup>4</sup> – the auditory imagery thus generated is retained for a second or two. This process can be iterated so that the auditory imagery remains in the phonological store for as long as the repetition continues (Baddeley 1992, p. 558). That is, one can produce an utterance in inner speech; the auditory imagery thus generated is retained long enough for one to produce the utterance again; and the process can be repeated indefinitely. The process is familiar: consider how you might try to remember, say, a phone number which you cannot immediately write down or record in your phone.

Next, inner speech need not and, we can assume, typically does not take the form of complete and grammatically correct sentences. Just as our external speech often takes the form of subsentential phrases, so does our inner speech. Most likely, inner speech is typically more truncated than external speech. This was Vygotsky’s view (Vygotsky 1934/1986, pp. 239-247). He believed, for example, that the object that an individual is speaking about in inner speech will always be apparent to them, so they need only express predicates about it

---

<sup>3</sup> I am grateful to Joanna Atkinson of the Deafness, Cognition and Language Research Centre at University College London for confirming that there is a lack of research on this.

One very interesting point is worth noting. It may be natural for hearing people to assume that the experience of inner sign would be primarily visual, just as the experience of inner speech is primarily auditory. However, things cannot be this simple. When a hearing person speaks externally, what they hear is basically the same as what others who are listening to them hear. However, when someone produces sign language externally, what they see is the reverse of what someone in front of them sees. The signer sees their signs from behind, as it were (Atkinson 2006, p. 706; the point is more explicit in Atkinson et al. 2007, p. 357). Accordingly, Atkinson (2006) speculates that ‘[f]uture research may reveal that mental representations of sign language are primarily kinaesthetic because muscular feedback predominates how an individual would experience their own productions’ (p. 706).

<sup>4</sup> Baddeley actually describes the articulatory control process as ‘somewhat analogous to inner speech’ but does not indicate what this might mean (1992, p. 558). Perhaps he had a somewhat curtailed notion of inner speech in mind. There does not seem to be anything about inner speech as it is being characterised in the text here which would prevent the auditory imagery involved in an utterance in inner speech from being retained in the phonological store.

(Vygotsky 1934/1986, pp. 243-244). In any event, it seems safe to say that, most of the time, the utterances we produce in inner speech are very truncated.<sup>5</sup>

One final feature can confidently be attributed to inner speech. We do not have the same control over our inner speech as we have over our external speech. Pathological cases aside, it is very easy to stop producing external speech; it is extremely difficult to quiet our inner speech.

We should also note one important respect in which we might expect to be able to make a confident claim about inner speech but cannot in fact do so. It is not clear whether producing inner speech necessarily involves subvocalisation in the sense of activity in the muscles involved in speaking externally. MacKay has written that the theory ‘that motor activity is necessary for perceiving internal speech and other forms of imagery has been extensively examined and conclusively rejected’ (1992, p. 133). By contrast, Stephane et al. claim that ‘it is generally agreed that subvocalizations accompany inner speech’ (it is clear from the text that the authors take subvocalisation to refer to activity in the speech muscles) (2001, p. 74). The matter is hardly settled by self-observation. It is plausible that producing inner speech does not involve movements of the speech muscles but, instead, involves faint kinaesthetic imagery of these movements (see e.g. Machery 2005 for some suggestions that inner speech involves kinaesthetic imagery; this source is discussed in Footnote 25). I suggest this finds no support in introspection. However, if one thinks inner speech does involve subvocalisation, it would be important to rule out this alternative hypothesis.<sup>6</sup>

---

<sup>5</sup> Martínez-Manrique & Vicente point to an example which nicely demonstrates just how truncated interpersonal speech often is (2010, p. 142). The example comes from Steven Pinker’s *The Language Instinct* (1994, pp. 222-224). Pinker provides some context: ‘in 1979 [Richard] Nixon, claiming to be acting on behalf of future historians, bugged his own office and began secretly taping all his conversations’ (1994, p. 224). As the Watergate scandal unfolded, the tapes were subpoenaed and transcribed. Pinker reproduces the transcript of a conversation between Nixon, his counsel, and his chief of staff, which was central to the Judiciary Committee of the House of Representatives recommending, in 1974, that Nixon be impeached. The transcribed conversation is basically incomprehensible. Pinker notes that this is partly because intonation etc. cannot be reproduced in a transcription and because the recording was of a low quality to start with. But the highly truncated phrasing has a lot to do with it too. ‘[E]ven when transcribed perfectly’, Pinker writes, ‘conversation is hard to interpret. People often speak in fragments, interrupting themselves in midsentence to reformulate the thought or change the subject’ (1994, p. 224).

Vygotsky also offers some interesting examples of abbreviation in the external use of language from Russian literature (1934/1986, pp. 237-242). In Tolstoy’s *Anna Karenina*, for instance, two characters, Kitty and Levin, are so aligned in their thinking as they declare their love for one another that they can communicate merely by writing the initial letters of words in the sentences they wish convey to each other (Vygotsky, 1934/1986, pp. 237-238, quoting from *Anna Karenina*, Part IV, Chapter 13).

There will be more on the abbreviation of inner speech in Section 2.3 of Chapter Two.

<sup>6</sup> We will see in Section 2.1 of the Appendix that the notion that inner speech does involve subvocalisation plays a role in an influential theory which seeks to explain auditory verbal hallucinations.



In terms of a positive characterisation, then, an inner speech utterance is a psychological experience involving the generation of a linguistic item. Inner speech is episodic. It involves auditory imagery and can be used to retain information in working memory. Inner speech utterances are typically subsentential. We do not have the same level of control over our inner speech as we do over our external speech. This much, I expect, can be said about inner speech without provoking debate, but it is as far as we can easily go. At this point, then, it makes sense to review some examples of the phenomenon.

### 3. Some Examples

There is, of course, a difficulty in trying to provide examples of inner speech: there is literally an infinity to choose from. Moreover, whatever examples one provides, there is the problem of organising them: there is no single, best taxonomy for such a diverse phenomenon. The best that can be done is to try to provide examples which will seem familiar; to show something of the variety of the phenomenon; and to organise the sample in some convenient and sensible way. What follows is intended to meet these criteria; it is not a thorough catalogue. Only relatively few of the examples which follow have the truncated nature which, I have just suggested, we can assume is typical of inner speech utterances. Efforts to capture this feature in print inevitably seem somewhat awkward and unconvincing.

First, there are instances of inner speech which seem to be directed towards solving problems. These instances of inner speech might take the form of questions which we seem to be asking ourselves, possibly followed by answers to them. For example, searching for lost keys: ‘Where did I have them last? Ah, in the kitchen.’ Or in the initial stages of thinking about an issue in political philosophy: ‘A right to vote is important. What do we need to establish this right?’ Such instances might also take the form of hypotheses which we seem to be suggesting to ourselves for consideration. A doctor considering why a patient’s sinus irritation is so persistent might produce the following inner speech utterance: ‘It could be an allergy.’ Or a soccer coach considering what to do with an ineffective defender: ‘Perhaps she’d be better on the wing.’

Next, there are instances of inner speech which seem to be directed towards making choices.<sup>7</sup> In thinking about what to have for lunch, one might produce the following series of inner speech utterances: ‘What should I have for lunch? I had sushi yesterday. I’ll just have a

---

<sup>7</sup> Of course, making a choice amounts to solving a certain kind of problem, i.e. a decision problem (thanks to Daniel Stoljar for this point) and there is certainly some similarity between the examples which are given for both categories. Still, there is an intuitive distinction between them which is worth honouring.

sandwich.’ When considering how to spread the investments in a retirement fund, one might, as it were, suggest to oneself: ‘Shares are too volatile. I’m going to put a bit more in property.’ Or when thinking about which candidate to support in an election: ‘The major parties are so cynical. I’m going to vote for an independent.’ Just a slight variation on this theme would be cases where one has already made a choice and produces an inner speech utterance expressing a reason for the decision, as if to rationalise it. For example, someone feeling some compunction about cancelling an engagement with one friend in order to see another might produce this inner speech utterance: ‘Well, I hadn’t seen her for much longer.’

Thirdly, there are inner speech utterances directed towards task-switching. While working on a paper but needing to attend to some errands, one might produce the utterance, ‘I’d better call the dentist now’, before doing just that. Or as a news program which one has been watching concludes: ‘Time to do the washing up.’ A subset of this category would be cases where one is performing a series of activities and produces inner speech utterances as one moves from each task to the next. While doing the different activities in one’s aerobics routine, for example, one might produce the utterances, ‘Three steps left, three steps right, jump.’ Or, while learning to drive a manual car and practising starting on a hill: ‘Clutch, gear, accelerator.’

Next, there are instances of inner speech associated with memory. We have already noted the role that inner speech can play in working memory. If one has to remember for a short period a phone number or a list of items to purchase at a shop, it can sometimes assist to recite the series of numerals or the list of items internally, perhaps over and over again. In these cases, inner speech seems to serve as a type of memory aid. Apart from the use of inner speech in working memory, there are also instances when one (perhaps suddenly) recalls a commitment or obligation or previously formed intention and produces an inner speech utterance expressing this. One might produce the inner speech utterance, ‘The bank!’, when recalling that they meant to deposit a cheque during their lunchbreak, or ‘The game!’, when realising that a sporting event which they intended to watch is due to start.

There are inner speech utterances which take the form of exhortations, imperatives or addresses directed to oneself. An anecdote which will be familiar to many philosophers provides an example of this. When thinking with a great deal of effort, Jerry Fodor apparently experiences inner speech utterances such as ‘C’mon, Jerry, you can do it!’<sup>8</sup> In a similar vein, an amateur runner nearing the end of a half-marathon might say in inner speech ‘You can do it’, or ‘Just two more kilometres’. Or a crossword enthusiast feeling less buoyant about her

---

<sup>8</sup> Or so the story goes. See, for example, Dennett 1991, p. 303. So far as I know, Fodor has never actually claimed this in print.

prospects for success at the day's puzzle might produce the utterance, 'Too hard. Time to give up.'

There are also inner speech utterances which take the form of exhortations, imperatives or addresses which we produce as if we were speaking aloud to someone else. This category would include utterances which one realises it may be better not to say aloud, temptation notwithstanding. Rushing along a crowded city footpath and nearly colliding with another, perhaps slightly inattentive pedestrian, one might fume, 'Get out of the way!' in inner, rather than external, speech. (Probably, many people are somewhat less well-mannered in their inner speech than in their external speech – and, indeed, somewhat less well-mannered than the example might suggest – given that we need not worry whether our inner speech conforms to social norms.) The category would also include utterances which one would like to, but cannot presently, express to someone else, perhaps on account of distance or circumstance. Watching one's daughter compete in a diving competition from the grandstand, one might anxiously produce the inner speech utterance, 'Don't forget to tuck your knees in!', as she prepares to jump from the platform. Or, while watching a friend giving a talk and observing that they are speaking too fast: 'Slow down! You need to slow down!' A variation on this theme would be addresses one directs to oneself at an earlier point in time. Regretting a bad decision, for example, one might produce the inner speech utterance, 'You're such a fool', as if actually speaking to the earlier version of oneself.

A similar category would comprise brief, exclamation-style utterances in inner speech. One might produce the inner speech utterance, 'What a pain!', when thinking of a difficult person they have just encountered; or 'Oops!' in reaction to some minor accident. One might produce the inner speech utterance, 'Oh, no!' on receiving very bad news; 'Yes!' upon learning of some good news; or 'Wow!' upon experiencing some surprise.

We sometimes (or perhaps often) make remarks in inner speech which are about the tasks and activities we are carrying out but which do not play any obvious role in facilitating these tasks or activities. Closing a door but realising that the latch has not engaged properly, one might produce the inner speech utterance, 'Ah, it hasn't quite shut', before reaching for the door handle again and applying a little more force. Preparing a cup of coffee, one might utter in inner speech, 'Now a teaspoon of sugar', before adding exactly this to the beverage. Walking along a footpath after some rain and seeing a deep-looking puddle, one might say in inner speech, 'Better steer clear of that'. In each of these cases, it seems very unlikely that the inner speech utterance plays any important role in making a decision or otherwise organising one's behaviour.

There are also instances of inner speech which do not serve any obvious purpose and do not seem to be associated with any task that one is carrying out. Just as I sometimes, in idle moments, softly recite in external speech a line once learnt in a high school English class, ‘I caught this morning morning’s minion’, so do I sometimes produce this utterance in inner speech. One could try to name in inner speech all the colours that one knows or to list Europe’s living monarchs. Maybe there is someone who enjoys repeating in inner speech the word, ‘palimpsest’.

#### 4. Vygotsky and the Development of Inner Speech

We have noted some features which can fairly safely be attributed to inner speech and we have reviewed some examples which will hopefully seem either familiar or plausible. We can take one further step in characterising the target phenomenon. This is to review Lev Vygotsky’s theory on the development of inner speech in children which is set out in his classic *Thought and Language* (1934/1986).<sup>9</sup> This theory will be relevant at a number of points throughout the dissertation, so it will serve to introduce it here.

For Vygotsky, ‘[t]he primary function of [external] speech, in both children and adults, is communication, social contact’ (1934/1986, p. 34). For this reason, Vygotsky believes, the first speech which children engage in – presumably something like communication with their parents – is ‘essentially social’ (1934/1986, p. 35). So how does *inner* speech develop? Like Piaget before him, Vygotsky observed that young children have a tendency sometimes to speak audibly about activities that they are carrying out (1934/1986, p. 25 ff.). Children engage in this practice, which Vygotsky (following Piaget) called ‘egocentric’ speech, until around the age of seven or eight (Vygotsky 1934/1986, p. 29, citing Piaget).<sup>10</sup> (Piaget referred to this type of speech as ‘egocentric speech’ primarily because the child engaging in it ‘does not attempt to place himself at the point of view of his hearer’ (1934/1986, p. 26, quoting Piaget 1959, p. 9)). Elaborating Piaget, Vygotsky writes that ‘[t]he child does not try to communicate,

---

<sup>9</sup> There are two English translations of this text which are typically relied upon now, Vygotsky 1934/1986 and Vygotsky 1934/1987 (in which the text is actually titled *Thinking and Speech*). It appears that both are deficient; see van der Veer & Yasnitsky 2011 for severe criticism of both. So far, I have been relying on the 1986 translation. I will continue to do so in this Introduction and in Chapter One. In Chapter Two, I deal at length with a theory developed by Charles Fernyhough. Fernyhough generally uses the 1987 translation. I will do so as well in that chapter.

<sup>10</sup> Vygotsky’s referencing is poor by modern standards. He does not provide a specific citation here, though he is generally discussing Piaget (1959) at this point in the text. (For a little on Vygotsky’s referencing, see the ‘Notes to References’ (Rieber & Carton 1987) at the end of a translation of another of Vygotsky’s texts. It is actually not at all clear to me whether this note is written by the editors of that volume, Robert W. Rieber and Aaron S. Carton, or the translators, Jane E. Knox and Carol B. Stevens. I have cited Rieber & Carton in the reference list.)

expects no answers, and often does not even care whether anyone listens to him' (1934/1986, p. 26.) Piaget thought that egocentric speech simply disappears at around age seven or eight (Vygotsky 1934/1986, pp. 29, 32).<sup>11</sup> Vygotsky, on the other hand, thought that it continues but that it is internalised – that children cease producing external utterances of this kind and replace them with utterances in inner speech (1934/1986, pp. 30-33). As he puts it: egocentric speech 'does not simply atrophy but "goes underground," i.e., turns into inner speech' (1934/1986, p. 33; see also pp. 86-88).<sup>12</sup>

Vygotsky points to a number of considerations in support of his conclusion. First, he points to what he takes to be two features which egocentric speech and inner speech share. He writes:

The first feature uniting the inner speech of adults with the egocentric speech of children is its function as speech-for-oneself. If one turns to Watson's experiment and asks a subject to solve some problem thinking aloud, one would find a striking similarity to the egocentric speech of children. (Vygotsky 1934/1986, p. 32).<sup>13</sup>

In addition, Vygotsky asserts, children's egocentric speech and adults' inner speech have the 'same structural characteristics' (1934/1986, p. 32). Utterances in either form of speech taken 'out of context ... would be incomprehensible to others because they omit to mention what is obvious to the speaker' (1934/1986, p. 32). Secondly, Vygotsky notes that, at the age when, it is hypothesised, egocentric speech is internalised, 'children facing difficult situations resort now to egocentric speech, now to silent reflection', surmising that 'the two can be functionally equivalent' (1934/1986, p. 33). Finally, he refers to '[o]bservations made by Auguste Lemaître and some other authors'<sup>14</sup> that 'the inner speech of schoolchildren is labile and unsettled', which he takes to suggest that inner speech at this age 'is a developmentally immature process that is still unstable and indefinite' (Vygotsky 1934/1986, p. 33; according to Alex Kozulin, the editor of Vygotsky 1934/1986 (p. 263, Endnote 11), Vygotsky is referring to Lemaître 1905).

There is a certain oddity about Vygotsky's theory. The notion that egocentric speech is internalised, 'goes underground', or 'turns into inner speech' is less straightforward than it might appear. If an activity which takes place in the external world becomes an activity which takes place in the mind, is this not a major transformation? Has the activity just 'go[ne] underground' and not actually been replaced with a different one?<sup>15</sup> Insofar as Vygotsky does not explain exactly how inner speech relates to the egocentric speech which precedes it

---

<sup>11</sup> Again, Vygotsky does not provide a specific citation here.

<sup>12</sup> Vygotsky's inverted commas here seem to be intended to bracket the informal expression rather than to indicate a quotation.

<sup>13</sup> Vygotsky is referring to John B. Watson but he does not cite a particular source.

<sup>14</sup> Vygotsky does not indicate who these 'other authors' are.

<sup>15</sup> Thanks to Daniel Stoljar for this point.

developmentally, his account is, we might say, somewhat incomplete.<sup>16</sup> Accordingly, it might be better to think of Vygotsky's contribution as being that he noticed that inner speech is a developmental successor to egocentric speech, whether or not it is an internalised version of egocentric speech.

\* \* \*

So, I have itemised some features which can be attributed to inner speech with a high degree of confidence; provided a range of examples of inner speech which will hopefully seem familiar or at least plausible; and reviewed Vygotsky's highly influential theory on the development of inner speech in individuals. The phenomenon should now be in focus. The way is clear to address our four questions about it.

---

<sup>16</sup> I will return to this point in Section 6.1 of Chapter One.

## Chapter One: Imagined Speech or Actual Speech?<sup>17</sup>

### 1. The Problem

Consider some mundane activities: going for a walk, making a salad, drinking a cup of coffee. What happens if you actually engage in one of these activities? You take strides at a regular pace, add ingredients to a bowl, sip from a cup. But it is quite different if you *imagine* engaging in one of these activities. No striding, adding ingredients, or sipping takes place. Rather than performing a series of actions in the external world, you somehow represent yourself performing the actions in your mind.

Is inner speech a type of actual speech? Or is it a type of imagined speech?<sup>18</sup> Generally, it is easy to say whether you are actually running, actually making a salad, actually writing a list, or just imagining doing one of these things. In the ordinary course of events, it would be very surprising if someone were to report, for example, that they were unsure whether they had just been going for a run or had just been imagining that they were going for a run<sup>19</sup> (see Mellor 1993, p. 11, for a very similar point). With inner speech, though, the situation is not so simple. If you produce an utterance in inner speech, you could reasonably think either that you are actually speaking or that you are just imagining speaking. You might think that you are actually speaking because you seem to be doing things like making assertions and asking questions, just as you do in external speech. Or you might think that you are just imagining speaking because the activity takes place wholly in your mind and does not involve producing words aloud.

My purpose in this chapter is to argue that inner speech is a type of actual speech: when one produces inner speech, one is actually speaking. In Section 2, I will set out what seem to be

---

<sup>17</sup> A significant amount of the material in this chapter has been published in the *Review of Philosophy and Psychology* (Gregory 2016).

<sup>18</sup> I am extremely grateful to Daniel Stoljar, who suggested this question to me, and to Inga Vermeulen, who independently raised with me a slightly narrower version of the same question. I should also here record a word of apology to Inga. It should have been noted in my paper in the *Review of Philosophy and Psychology* that Inga, as well as Daniel, had raised this issue with me, but I had forgotten that she had done so, and did not recall this until revisiting some old notes after the paper was published. This will be reiterated in any future published work on the topic.

<sup>19</sup> Setting aside individuals suffering from some pathology – and, of course, philosophers. Regarding the latter, one is reminded here of the famous dreaming passage from the work of Zhuangzi, the ancient Chinese philosopher. Zhuangzi reports that he dreamed he was a butterfly. Then he woke, and could not be sure if he was a person who had just dreamed he was a butterfly, or a butterfly now dreaming that he was a person. Accordingly, he did not know whether he was dreaming at any point during the episode (Chuang-tzū, 1981). If dreaming is a kind of imagining, which seems plausible, then we can say that Zhuangzi did not know whether he had actually been doing something – say, fluttering about – or just imagining doing so.

the best possible versions of the views that inner speech is a type of imagined speech and that inner speech is a type of actual speech. I will also consider the possibility that these characterisations do not cover the field and that there may be some other characterisation which should be preferred to both of them. In Section 3, I will present a number of arguments to the effect that inner speech is a type of actual speech, not a type of imagined speech. In Section 4, I will respond to a possible objection. In Section 5, I will calibrate the strength of the case for the view that inner speech is a type of actual speech and, in Section 6, I will explore some of the implications of this view.

## **2. The Possible Views**

Before attempting to adjudicate whether inner speech is a kind of imagined speech or a kind of actual speech, it is necessary to have in mind the form that each of the two theories might take. Accordingly, I will provide a sketch in this section of what I will call the ‘Imagined Speech View’ and the ‘Actual Speech View’. I will also note some considerations which provide initial support for each of the views, so that we can see that there is genuinely an issue to be settled. Then I will consider the possibility that inner speech might best be characterised as neither imagined speech nor actual speech, but as a distinct phenomenon which has some, but not all, of the features of speech. I will call this view the ‘Quasi-Speech View’ and show that it is really just a variation on the Actual Speech View – though I will return to it again towards the end of the chapter.

It may be that there are slightly different ways of developing the view that inner speech is imagined speech. Accordingly, one might think that the Imagined Speech View could be improved upon. If this is the case, then there could be a concern that the arguments made against the Imagined Speech View in Section 3 would not have force against an alternative, better version of the view. In Section 5, however, it will become clear that the commitments of the Imagined Speech View which the arguments against it exploit are commitments which no reasonable version of the view could avoid. Insofar as those commitments are unavoidable, it will not matter if one thinks the Imagined Speech View could be improved upon in other ways. There is little room for disagreement with respect to the Actual Speech View which, as we will see, basically holds that inner speech is just silent speech.

In setting out the different views, I am going to proceed on the assumption that external speech is the paradigm of actual speech. This seems entirely reasonable: while there might be disagreement as to whether inner speech is a type of actual speech, there can be no such disagreement about external speech.



## 2.1. The Imagined Speech View

There are some suggestions in the literature that inner speech is imagined speech or that inner speech is a type of imagined speech (e.g. Machery 2005; Langland-Hassan 2008; Roessler 2013), but the idea is not developed in detail. In setting out the Imagined Speech View, I will draw on some of these suggestions, and also seek to develop the idea somewhat further.

First, there is a familiar distinction between what we can call, following McGinn (2009), ‘sensory’ imagination and ‘cognitive’ imagination (p. 595). McGinn describes this distinction in approximate terms as the distinction ‘between forming a mental image and entertaining a possibility in the conceptual style’ (p. 595). He offers these examples:

I may form a visual image of the colour blue or of a friend's face, on the one hand, and I may consider the possibility that a skyscraper might at some future time be built on this very spot, on the other. (p. 595).

McGinn’s examples of sensory imagination involve the visual modality: imagining a colour or a familiar face. But sensory imagination can involve other modalities including – importantly for present purposes – the auditory modality. One can imagine the chirping of a canary or the screech of a car tyre or the beginning of the Moonlight Sonata. As a preliminary point, the Imagined Speech View urges that inner speech is sensorily imagined speech, not cognitively imagined speech. One can conceive of the possibility of speaking, but this has nothing to do with inner speech.

Now, it is obvious that there can be instances of sensorily imagined speech, but the Imagined Speech View does not hold that every instance of sensorily imagined speech is an instance of inner speech. There are two kinds of cases to note here. First, one can just imagine some linguistic item (e.g. words, a sentence) being spoken in some generic human voice.<sup>20</sup> Secondly, one can imagine the sound of particular others speaking. You might, for example, imagine the sound of a friend speaking, or of two friends in conversation. Instances of these two kinds of cases involve sensorily imagined speech, but they are not instances of inner speech, as ‘inner speech’ is usually understood.<sup>21</sup>

---

<sup>20</sup> Thanks to Daniel Stoljar for a suggestion leading to this point.

<sup>21</sup> Remarkably, there is some empirical evidence which has been taken to show that inner speech can involve the voices of others. The relevant studies are McCarthy-Jones & Fernyhough (2011) and Hurlburt et al. (2013). I will discuss these in Section 4.2 of the Appendix. In a word, though, McCarthy-Jones & Fernyhough rather uncarefully treat all conscious linguistic episodes as instances of inner speech, and the evidence reported by Hurlburt et al. is so weak that it should simply be set aside. But this does not matter for present purposes. The key point is simply that, on any ordinary understanding of inner speech, inner speech does not just involve imagining hearing some linguistic item being spoken, either in some generic human voice or in the voice of a particular other or others.

So it is clear already that the Imagined Speech View holds that inner speech is a particular kind of sensorily imagined speech. But what, on the Imagined Speech View, distinguishes inner speech from other kinds of sensorily imagined speech? This is the key issue for the Imagined Speech View, so it deserves careful attention.

Having noted that an episode of imagining someone else, or some group of others, speaking will not be an episode of inner speech, one might be tempted to think that all and only instances of imagining oneself speaking are instances of inner speech. But this is too simple. If I were to imagine seeing and hearing myself give the State of the Union Address before a House of Representatives full of dignitaries, with my voice amplified and reverberating around the chamber, this would involve imagining myself speaking, but it would also not be an episode of what we ordinarily think of as inner speech.

Another possibility is that inner speech involves imagining performing the act of speaking. Roessler is very explicit in suggesting this: ‘to identify an experience as an act of inner speech you would need to recognize it ... as one of imagining a *performance* rather than an *observation* of a speech act’ (2013, p. 662, emphasis original). And Langland-Hassan raises the possibility that ‘inner speech is a kind of imagination or *mere* simulation of action (in this case, simulation of speaking)’ (2008, p. 386, emphasis original).

The Imagined Speech View does hold that producing inner speech involves imagining performing the act of speaking, as opposed to observing it, but this is still not enough to separate inner speech from other episodes of imagining oneself speaking. To see this, we need only revise the example above. Suppose I imagine *performing the act* of giving the State of the Union Address before a House of Representatives full of dignitaries, with my voice amplified and reverberating around the chamber. In this case, I would be imagining performing an action or a series of actions, but this still does not seem like inner speech.

The comparison just made between imagining observing oneself giving a speech as opposed to imagining giving a speech recalls a distinction which is common in philosophical discussions of imagination. This is the distinction between (what is sometimes called) imagining ‘from the inside’ and the ‘outside’. An early example of this distinction can be found in Williams (1973). Imagining from the inside can be taken to involve imagining from an internal point of view; imagining from the outside involves imagining from an external point of view.<sup>22</sup> We can get clearer on the distinction with another example. Take the case of imagining playing the guitar. If I imagine playing the guitar from the inside (i.e. from an internal point of view), it will be as if I can look down and see my hands strumming the strings

---

<sup>22</sup> This is not a new way to explicate the difference; it is just one intuitive approach.

of the guitar, look out and perhaps see some people who are listening. By contrast, if I imagine playing the guitar from the outside (i.e. from an external point of view), it will be as if I can see myself from without, perhaps sitting and strumming the guitar.

A proponent of the Imagined Speech View might try to recruit this distinction to separate inner speech from other episodes of imagining oneself speaking. When one engages in inner speech, one does not seem to see or otherwise have an experience as of oneself speaking from a point of view outside one's own body, as would need to be the case if inner speech involved imagining speaking from the outside. So it might be claimed that inner speech involves imagining speaking from the inside, whereas other episodes of imagining oneself speaking involve imagining speaking from the outside.

It is fairly easy to see, though, that imagining speaking from the inside (as opposed to the outside) is the same thing as imagining performing the action of speaking (as opposed to imagining observing it). If one imagines performing an action, one can only do so from the inside; if one imagines observing an action, one can only do so from the outside. Accordingly, the suggestion that producing inner speech involves imagining speaking from the inside is subject to the same counterexample as the suggestion that producing inner speech involves imagining performing the action of speaking. Imagining giving the State of the Union Address before a House of Representatives full of dignitaries, with my voice amplified and reverberating around the chamber, involves imagining from the inside. But, as noted above, an episode such as this is not what we would think of as an episode of inner speech.

One final possibility is that inner speech involves imagining speaking in a way which has a very bare or basic phenomenology. The imagery that it involves is just auditory imagery of word-sounds and, perhaps, kinaesthetic imagery of the movements involved in speaking aloud.<sup>23</sup> One finds suggestions along these lines in the literature. Langland-Hassan, for example, writes that 'the sensory character of inner speech is not nearly as rich as that of actual speech perception' (2008, p. 293). And Martínez-Manrique & Vicente write that '[t]he Vygotskian approach to inner speech regards it as a phenomenologically degraded form of our own talking ... [i]n this view, inner speech ... lacks pitch and volume' (2010, p. 142).<sup>24</sup> (Let us set aside the suggestion that inner speech is a 'form of our own talking', wording which is more compatible with the view that inner speech is a type of actual speech. The key point here is just the idea that inner speech is 'phenomenologically degraded'.)

---

<sup>23</sup> See Section 2 of the Introduction to the dissertation for some brief discussion of the possibility that inner speech involves kinaesthetic imagery, as well as Footnote 25, dealing with Machery 2005.

<sup>24</sup> There are echoes in these quotations of the Humean notion of imagination, according to which sensory imagery involves fainter representations than actual perceptual experience. Thanks to Daniel Stoljar for this.

There is merit to this suggestion that the sensory phenomenology of inner speech is, in some sense, bare or basic. Certainly, it excludes cases such as imagining performing the act of giving the State of the Union Address before a House of Representatives full of dignitaries, which seem too elaborate to be designated instances of inner speech. However, there are a couple of complications with this suggestion also.

First, the suggestion needs to be made a little more precise. The idea that the auditory imagery involved in inner speech lacks, for example, pitch and volume, seems extremely dubious. Surely, one cannot imagine a sound which lacks pitch and volume any more than one can imagine a patch of colour which lacks hue and saturation.

This first problem can be addressed. Rather than holding that producing inner speech involves producing imagery which lacks one or more of the characteristics which we expect imagery in a particular modality to have, the Imagined Speech View holds that producing inner speech involves producing imagery in only the auditory modality (as well as, possibly, producing some faint kinaesthetic imagery). This allows that inner speech seems to have qualities such as pitch, which seems unavoidable, but still provides that inner speech has a bare phenomenology insofar as it does not involve, for example, visual imagery.

The more serious problem is just that inner speech also seems to have other features. The suggestion that inner speech has a bare sensory phenomenology is not even close to a complete characterisation of inner speech. Supposing inner speech is a type of imagined speech, it does seem to involve imagining oneself speaking rather than imagining some other person or other people speaking. It does seem to involve performing (rather than observing) the act of speaking. And, to say the same thing with a different emphasis, it does seem to involve imagining speaking from the inside rather than from the outside.

What the Imagined Speech View claims, then, is that inner speech has all of these features. Any instance of imagining oneself speaking which has all of these features is an instance of inner speech. On the face of it, this certainly seems plausible.<sup>25</sup>

---

<sup>25</sup> At different points, Edouard Machery (2005) uses terminology which is consistent with a number of different possible characterisations of inner speech as a type of imagined speech. Sometimes, his language is strongly suggestive of a view that producing inner speech does not in fact involve imagining speaking but just involves imagining an utterance, whatever exactly this might involve – phrasing such as ‘imagine a sentence’ (2005, p. 478), ‘imagined sentences’ (2005, p. 480), and ‘imagining a sentence’ (2005, p. 482). But he also speaks of ‘articulatory’ (as well as auditory) imagery, so he presumably thinks that inner speech does involve imagining speaking (2005, pp. 471, 476, 477, 483). And he cites empirical evidence (Shergill et al. 2001) which he takes to support the notion ‘that inner speech is a form of verbal imagination: I imagine hearing a sentence that I attribute to myself’ (Machery 2005 p. 477; footnote omitted). As far as I can tell, his view is that producing inner speech involves imagining speaking; when he suggests that inner speech involves imagined utterances, he means that we imagine producing these utterances. However, given the lack of clarity, I have not considered Machery’s paper in the text.

It is possible to make one further observation. The Imagined Speech View treats inner speech as possible speech.<sup>26</sup> Insofar as the Imagined Speech View takes inner speech to involve, *inter alia*, performing the action of speaking, it takes inner speech to involve imagining performing an action that one could, at least in principle, actually perform, i.e. by producing external speech. The ‘in principle’ restriction is important here. It may be that, as a matter of fact, someone can produce utterances in inner speech but cannot speak aloud, perhaps on account of some injury to the larynx. The Imagined Speech View would emphasise that this is a practical barrier to producing actual, external speech and that, were it not for the injury, the individual would be able to produce the actual, external speech which their inner speech resembles.

This is how the Imagined Speech View characterises inner speech. How might the view be motivated? The key consideration is that there are some features of external speech – which, again, we can think of as the paradigm of actual speech – which are plainly lacking from inner speech. These are features which one could plausibly take to be essential to actual speech, so their absence will potentially be damaging to the view that inner speech is a form of actual speech, no matter how exactly that view might be developed. However, they are not features that we would expect inner speech to instantiate if inner speech is a kind of imagined speech.

There are three such features which seem salient. The first is simply that external speech involves vocalisation<sup>27</sup> and, consequently, the actual production of sounds, whereas inner speech does not. The second is that external speech can be used in the service of communication, whereas inner speech cannot. The third is a little more complicated and requires some explanation.

There are at least some actions which one can perform by producing an utterance of external speech in certain circumstances which one cannot perform by producing an utterance of inner speech. Take the action of a judge in sentencing a convict. In appropriate circumstances, a judge can issue a sentence by uttering in external speech something like ‘I sentence you to ...’. Plainly, producing a corresponding utterance in inner speech even in exactly the same circumstances will not be sufficient to perform the action of sentencing. The same may apply to other actions such as betting, marrying, apologising – roughly, to the category of actions

---

<sup>26</sup> Thanks to David Chalmers for this point.

<sup>27</sup> This was suggested to me by an audience member in a presentation at the Serious Metaphysics Group at the University of Cambridge in April 2014. There will be further discussion of this point in Section 4.

which can be carried out through what J. L. Austin calls ‘performative utterances’ (Austin 1962).<sup>28</sup>

Here, then, are three features of external speech – it is vocalised, it facilitates communication, and it enables performative utterances – which are absent from inner speech. If one takes any of these features to be essential to actual speech, then one will have a reason to deny that inner speech is a kind of actual speech. Moreover, they are features which we would not expect to find in inner speech on the hypothesis that inner speech is a type of imagined speech.

This is how the Imagined Speech View proceeds. It offers a plausible characterisation of inner speech as imagined speech. It then seeks to exclude the rival view by pointing to potentially important features of external speech which are not instantiated by inner speech.

## 2.2. The Actual Speech View

The Actual Speech View claims that there are two types of speech which can with equal legitimacy be designated ‘actual speech’: one is external speech; the other is inner speech. The view has historical forerunners. It seems to be how Ryle, for example, conceives of inner speech, referring to our ‘silent monologues’ (Ryle 1949, p. 184). An example of a more recent suggestion to the effect that inner speech is a kind of actual speech comes from Philip Gerrans, although Gerrans actually makes his point using the language of imagination. He writes that ‘inner speech is a form of imaginary action’ (Gerrans 2015, p. 296). He then explains in a footnote that ‘[e]quating inner speech with *imaginary action* limits the type of imagination involved to that described by Jeannerod, 2006: the covert rehearsal of an action without an overt component’ (Gerrans, 2015 p. 296; emphasis original). The idea that inner speech is covert speech is really what the actual speech view holds.

In a nutshell, the Actual Speech View acknowledges, but seeks to de-emphasise, the differences between external speech and inner speech, and highlights the similarities. First, the view holds that inner speech need not instantiate every feature of external speech in order to be considered a kind of actual speech. Two of the features which are emphasised by the Imagined Speech View as lacking from inner speech are only contingently lacking from it, and the other seems considerably less important in light of this fact. Accordingly, the absence

---

<sup>28</sup> I say *roughly* the category of what Austin calls ‘performative utterances’ because there may be room for disagreement as to whether some items in that category can be performed in inner speech. For example, Austin takes promising to be a performative utterance, and it has been suggested to me that one can make a promise to oneself in inner speech. This seems dubious to me – for who exactly would have a grievance if the purported promise were not fulfilled? – but perhaps there is no single correct view here.

of these features does not provide a reason to believe that inner speech is not a type of actual speech. To see this, let us briefly work through the features of external speech which, as we just saw, a proponent of the Imagined Speech View might highlight, though in a slightly different order.

It is true that external speech facilitates communication, whereas inner speech does not. But it could be otherwise. It is only contingently true that we need to produce sounds or inscriptions etc. in order to communicate with one another. If we somehow had unmediated access to one another's inner speech – and there is nothing a priori inconceivable about this kind of telepathy – then inner speech would be facilitative of communication. It is not that inner speech per se cannot facilitate communication. The same applies to the point that external speech is enabling of performative utterances, whereas inner speech is not. If we had access to one another's inner speech, then it would be enabling of performative utterances. A judge *could* sentence a convict by producing an appropriate utterance in inner speech. Finally, if we were able to communicate with one another and produce performative utterances without vocalising our speech, then it might be doubted that vocalisation is an essential feature of speech. For everything that we do with external speech, we could also do with internal speech.<sup>29</sup>

Next, a proponent of the Actual Speech View can emphasise three striking, and closely connected, similarities between inner speech and external speech which, again, we can take as the paradigm of actual speech. The first is a simple point about the phenomenology of both inner speech and external speech: producing inner speech feels like speaking (albeit silently), not like imagining speaking. We produce inner speech in our minds, but we do not feel as if we are merely simulating speaking, to borrow Langland-Hassan's terminology. At present, this observation is very impressionistic, but the next point, as well as being independently noteworthy, will also provide some support for it.

The second point is this: as noted in the first section of this chapter, we can produce utterances in inner speech which seem to be of the same kind as many of the utterances we produce in external speech. Most obviously, a great many – perhaps most – of our utterances in inner speech seem to take the form of assertions and questions. This requires some elaboration, as does its connection to the first point.

As we saw, the Imagined Speech View points out that we cannot produce (at least many kinds of) performative utterances in inner speech. It has some justification for this: a performative

---

<sup>29</sup> This is not to say that speech must or should be characterised by the functions that it serves; it is just a position that a proponent of the Actual Speech View could take. In Section 4, there will be considerable discussion of the possibility that it is the generation of a concrete token instantiating a linguistic type which is really essential to speech (i.e. something about the nature of speech itself), rather than any function or functions that we think of speech as performing.

utterance requires a particular context which is not present in inner speech. For example, a judge's utterance, 'I sentence you to ...', will only be effective in sentencing a convict if delivered in a courtroom after a verdict has been returned. But one might think that the same is true of assertions and questions. For is a certain context not also necessary to make assertions or ask questions? Specifically, must there not be a listener, to whom one makes an assertion or of whom one asks a question? And if there must be a listener in order for a speaker to do these things, then is not the claim of the Actual Speech View that one can do them in inner speech just implausible?

There is a difference between performative utterances on the one hand, and assertions and questions on the other, that the Actual Speech View can appeal to in addressing these questions. Absent the appropriate context, (purported) performative utterances are completely incongruous. Austin writes the following of the performative utterance of marrying: '[w]hen I say, before the registrar or altar, &c., "I do", I am not reporting on a marriage: I am indulging in it' (1962, p. 6). Of course, one indulges (as Austin puts it) in no marriage if one produces exactly the same utterance in external speech but not in the appropriate circumstances. But the utterance would also be deficient in another way: it would simply not make sense. A bystander who overheard the utterance (and who understood the words and their significance in wedding ceremonies etc.) would wonder what the speaker was doing, given that the circumstances did not permit of the possibility that the speaker was marrying someone. Yet assertions made to ourselves, and questions asked of ourselves, do make sense. A bystander who overheard someone making an assertion to him/herself or asking a question of him/herself in external speech would not wonder what the speaker was doing. In a nutshell, then, the Actual Speech View holds that the notion of producing a performative utterance absent the appropriate context is just incoherent, whereas the notion of making an assertion or asking a question in the absence of a listener is not. In fact, it is very natural to say that we sometimes do precisely that.

The observation that it seems that many of the utterances that we produce in inner speech are of the same kinds as many of the utterances that we produce in external speech helps us to see why producing inner speech feels like actually speaking, not like imagining speaking. We feel as if we are speaking because we are (at least very often) actually producing assertions, questions etc.

The next similarity between inner speech and external speech that the Actual Speech View emphasises relates to assertions in particular. Assertions made in inner speech and assertions made in external speech are both evaluable as either true or false.<sup>30</sup> Of course, no one other

---

<sup>30</sup> Or, strictly, they express propositions which are evaluable as either true or false.



than the individual who produces an assertion in inner speech is in a position to determine whether the assertion is true or false, and it may be that ignorance prevents the relevant individual from making this determination as well. Still, the assertion itself has the property of being true or false, and the individual who produces the assertion *can* determine this, if they are inclined to do so and have, or acquire, the necessary information. (In many instances, of course, it will be entirely obvious to a subject whether something that they say in inner speech is true or false.) It might be objected that imagined assertions can also be true or false. This is so, but the point can be developed slightly to avoid this objection. An individual who produces an assertion in either external speech or inner speech is right or wrong – they speak truly or falsely – but this is not the case if one imagines producing an assertion. One can imagine producing an assertion which is false, but one is not in any way wrong if one does so.<sup>31</sup>

These similarities noted, the Actual Speech View points out that it can also account for the features of inner speech which the Imagined Speech View seemed well-equipped to explain. Supposing that producing inner speech amounts to producing actual speech, it makes sense that producing inner speech involves oneself speaking, rather than another person or other people. Supposing that producing inner speech amounts to producing actual speech, it makes sense that producing inner speech involves performing (rather than observing) the act of speaking. And supposing that producing inner speech amounts to producing actual speech, it makes sense that we seem to perform this act from our own, internal point of view, as it were. After all, there is no way that one can perform an act from an external point of view.

Finally, it is important to note that the Actual Speech View accepts that producing inner speech involves producing mental imagery. A view that did not accept this would be obviously false. What the Actual Speech View claims is that there is a kind of actual speech which consists of auditory images, rather than sounds. It holds, that is, that there is a kind of actual speech that takes place wholly in the mind.

This is how the Actual Speech View proceeds. It accepts the fact that inner speech lacks certain characteristics of external speech but urges that these characteristics are not essential to actual speech; it emphasises similarities between inner speech and external speech; and it urges that we can speak in mental imagery, just as much as we can speak aloud.

---

<sup>31</sup> Thanks to Daniel Stoljar for the suggestions in this paragraph.

### 2.3. Quasi-Speech

It might be suggested that the theory that inner speech is a type of imagined speech and the theory that inner speech is a type of actual speech, however they are developed, do not exhaust the possible accounts of the nature of inner speech. In particular, one might think that inner speech has some, but not all, of the features of actual speech; that inner speech is not imagined speech, but is still not quite the same thing as actual speech. Inner speech, one might say, is best thought of as a sort of quasi-speech. Let us call this view the Quasi-Speech View.

In making the case for the Quasi-Speech View, one could recruit some of the considerations already appealed to in developing the Actual Speech View and the Imagined Speech View. It is true that producing inner speech feels like actually speaking; that one can produce in inner speech at least some of the kinds of utterances that one can produce in external speech; and that assertions produced in inner speech are truth-evaluable just like assertions produced in external speech. But it is also true that inner speech is not vocalised and that it does not, as a matter of fact, facilitate communication or enable performative utterances. If one is impressed by both the similarities and the differences between inner speech and external speech, then it would be natural to believe that inner speech is almost, but still not quite, actual speech.

On the face of it, though, the Quasi-Speech View does not differ in any crucial way from the Actual Speech View. The Actual Speech View does not hold that there are no differences between inner speech and external speech; it explicitly acknowledges this. The view would deny the obvious if it denied that there are *any* differences between inner speech and external speech.

Is there a way that the Quasi-Speech View can be developed so that it is distinct from the Actual Speech View? There may be, by elaborating one of the considerations that motivates the Imagined Speech View. It is natural to think that the production of a linguistic item always involves the generation of some concrete, physical token. When one speaks externally, one generates a series of sounds. This is just a consequence of the fact that producing external speech involves vocalising. Similarly, when one writes, one produces print, either handwritten or typed, or a series of cells in Braille. So, if producing inner speech involves actually producing linguistic items, then we should expect that it would involve producing concrete, physical tokens of some kind, which does not seem to be the case.

In Section 4, I will consider whether this point can be appealed to in order to separate the Quasi-Speech View from the Actual Speech View, and also explore whether it can be recruited in making an argument against the Actual Speech View and for the Imagined Speech View. I will show, however, that there is no strong motivation for the point, and that it cannot be convincingly developed in either such way. Until then, I am going to proceed on the

assumption that the Imagined Speech View and the Actual Speech View do cover the field of possible characterisations of inner speech. No other features of external speech, which inner speech lacks and which might be taken to be essential to actual speech, seem to suggest themselves. And, if the view that inner speech has some, but not all, of the features of actual speech ultimately collapses into the Actual Speech View, then it is hard to see what positions are available other than the theory that inner speech is a kind of imagined speech, and the theory that inner speech is a kind of actual speech.<sup>32</sup>

### **3. Inner Speech is Actual Speech**

There are basically two views, then, about the nature of inner speech, both of which seem at least initially plausible. Why think that the Actual Speech View is right? In this section, I will present four arguments for this conclusion. These include both arguments directly supporting the Actual Speech View and arguments against the Imagined Speech View. Insofar as the Actual Speech View and the Imagined Speech View are being treated as exhaustive of the possible characterisations of inner speech, the arguments against the Imagined Speech View also, of course, provide support for the Actual Speech View.

Two caveats should be noted before beginning. The first is the point from the end of the last section, that the possibility of separating the Quasi-Speech View from the Actual Speech View will be revisited later in the chapter. The second is also a point foreshadowed earlier, that there may be slightly different ways that one could develop the view that inner speech is a kind of imagined speech. Again, I will also revisit this point later in the chapter, and show that the arguments against the Imagined Speech View which are offered in this section depend only upon commitments which no reasonable version of the view could avoid.

#### **3.1. The Embedding Argument**

The Embedding Argument turns on a feature of inner speech which seems to be inconsistent with the Imagined Speech View, but which is wholly compatible with the Actual Speech View. Roughly, this is that one can imagine producing inner speech. Making this argument will involve a number of steps. First, I will present the argument in slightly stronger terms than can ultimately be sustained. I will then consider a convincing objection to the argument

---

<sup>32</sup> Thanks to David Chalmers for suggesting the possibility of a Quasi-Speech View, and for suggesting that such a view might simply collapse into the Actual Speech View.

and reformulate it in such a way that it is not subject to this objection. Finally, I will consider some lingering concerns that one might have.

### **3.1.1. Stating the Argument**

As we saw, on the Imagined Speech View, producing inner speech involves imagining speaking from the inside (or, equivalently, it involves imagining performing the action of speaking).

Now, consider the following argument:

1. It is not possible<sup>33</sup> to imagine from the inside imagining from the inside.<sup>34</sup>
2. It is not possible to imagine from the inside imagining speaking from the inside (a fortiori from 1).
3. It is possible to imagine from the inside engaging in inner speech.
4. Inner speech is not imagining speaking from the inside (from 2 and 3).

Let us see what support the premises have.

### **3.1.2. Premises 1 and 2**

Premise 1 is that it is not possible to imagine from the inside imagining from the inside. To begin, recall how the notion of imagining from the inside is explicated in the Imagined Speech View. To imagine doing something from the inside is to imagine doing it from an internal point of view. Reiterating the example from Section 2.1 above, if I imagine playing a guitar from the inside (i.e. from an internal point of view), it will be as if I can look down and see my hands strumming at the guitar, look out and perhaps see some people who are listening. Now, notice the following: once one is imagining doing something from an internal point of view, one cannot imagine taking up that point of view again. Once I am imagining doing something from the inside, I cannot, as it were, go further inside.

One way to see this is to think of what taking up a point of view in the external world involves. Once I am in the position of the guitar player – that is, once I am actually playing the guitar, so that I can look down and see my hands strumming, look out and see an audience etc. – then I have that point of view. But a point of view is final: there is nowhere I can go to adopt this

---

<sup>33</sup> The possibility statements in 1-3 should be read as having narrow scope. Thanks to Daniel Stoljar for this observation.

<sup>34</sup> Zeno Vendler (1979, p. 166) also makes this claim, and D. H. Mellor (1993, p. 12) makes a very similar one.

point of view again, for I already have it. The same applies with imagination. Once I imagine having the point of view of the guitar player, there is nowhere else I can imagine going to adopt the same point of view, for I am already imagining having that point of view.

Premise 2 is that it is not possible to imagine from the inside imagining speaking from the inside. This is just an application of Premise 1.

### **3.1.3. Premise 3**

Premise 3 is that it is possible to imagine from the inside engaging in inner speech. We can demonstrate the truth of this simply by constructing an example. Imagine from the inside that you are on a putting green on a golf course, preparing to putt. Surveying the green, you produce the inner speech utterance, 'Looks like there's a little slope from the left to the right'. This episode of imagination from the inside seems easy to generate, including producing the inner speech utterance which, we might say, is embedded in the episode. The reason is that engaging in inner speech does not require adopting a new point of view. Like speaking externally or playing the guitar, it is something you do from the point of view which you occupy or imagine occupying. There is no problem, then, of imagining taking up some point of view and then of having somehow to imagine taking it up again.

### **3.1.4. A Refinement**

An objection to Premises 1 and 2 could take the form of counter-example. Consider the following case. An individual is imagining from the inside observing a beautiful scene in nature, perhaps the view over a valley from a certain vantage point. Within that context, they imagine thinking about how they might describe the scene to a friend. In doing so, they begin to imagine from the inside later describing the scene to that friend. That is, they imagine from the inside being in a particular situation – perhaps in a favourite cafe – speaking to their friend. This does seem like a case where one episode of imagining from the inside (describing the scene to a friend) takes place within another episode of imagining from the inside (observing the beautiful scene). Accordingly, it brings into question the truth of premises 1 and 2 in the argument above.<sup>35</sup>

Fortunately, the argument can be modified to avoid responses of this kind – and we already have the resources to make the necessary modification. What makes the case just described possible is that it involves two separate points of view. The subject adopts one point of view

---

<sup>35</sup> Thanks to Alma Barner for this point (I have modified her example).

in imagining from the inside observing the beautiful scene. Within that episode of imagination, the subject then imagines adopting a second point of view – specifically, at a later point in time – speaking to a friend.<sup>36</sup> What is not possible is to imagine from the inside imagining from the inside *from the same point of view*.

This is the contrast which is important for present purposes. For consider: in the scenario described, the individual *could* imagine from the inside observing the beautiful scene and producing some inner speech utterances about it (e.g. ‘This is so beautiful’). As above, this is possible precisely because imagining engaging in inner speech from the inside does not involve imagining taking up a second point of view. Inner speech is something you engage in from whatever point of view you occupy or imagine occupying. But our subject could not imagine from the inside imagining the beautiful scene from the inside (without a change of point of view). And they could not imagine from the inside imagining speaking about the beautiful scene from the inside (without a change of point of view).

To incorporate this point, the argument can be restated as follows:

1. It is not possible to imagine from the inside imagining from the inside (without a change of point of view).
2. It is not possible to imagine from the inside imagining speaking from the inside (without a change of point of view) (a fortiori from 1).
3. It is possible to imagine from the inside engaging in inner speech (without a change of point of view).
4. Inner speech is not imagining speaking from the inside (from 2 and 3).

### 3.1.5. Two Lingering Concerns

There are two lingering concerns that one might have about the Embedding Argument. Both of these concerns relate to Premise 1 and, by extension, to Premise 2. First, it might be suggested that it is possible to imagine from the inside imagining performing some action from the inside (without a change of point of view), but that such an episode will simply collapse into a single instance of imagining performing the action from the inside. That is, it might be suggested that there will be a case of what Shaun Nichols describes as ‘collapsed iteration’ (2003, p. 182). Gregory Currie explains this phenomenon as follows: ‘Try imagining someone else imagining something; what tends to happen is that you end up imagining it yourself’ (1995, p. 161). Perhaps what happens when you try to imagine from the inside

---

<sup>36</sup> Shaun Nichols (2003) gives an account of cases like this along the same lines.

imagining speaking from the inside is that you ‘end up’ just imagining speaking from the inside.<sup>37</sup>

There is a significant difference, however, between the notion of collapsed iteration and the notion of imagining from the inside imagining performing some action from the inside (without a change in point of view). There is nothing incoherent about imagining someone else imagining something, even if what ‘tends to happen [when you try] is that you end up imagining it yourself’. By contrast, on the understanding of imagining from the inside which I have specified, it is not even possible in principle to imagine from the inside imagining performing some action from the inside (without a change of point of view). Once you have taken up a particular point of view, there is simply no way of taking it up again. Yet it is possible – in principle and in practice – to imagine from the inside engaging in inner speech (without a change of point of view).

Secondly, it has been suggested to me that it is possible to imagine from the inside imagining performing some action from the inside (without a change in point of view), but that it just involves more ‘distancing’ than imagining performing some action from the inside.<sup>38</sup> What might this ‘distancing’ involve? It seems to be a phenomenological notion: one might feel that one is somehow further removed from an imagined action which takes place within another episode of imagination. And one can, I think, at least get a sense of the relevant phenomenology by considering the example from the last subsection. Recall the individual who imagines from the inside observing a beautiful scene in nature and who, within that context, imagines from the inside describing the scene to a friend. It does seem natural to say that the individual may feel further removed, in some sense, from the imagined describing of the scene than from the imagined observing of it. (Interestingly, this concern seems to move in the opposite direction, as it were, from the first, which was that imagining from the inside imagining from the inside collapses into just imagining from the inside.)

The difficulty with the suggestion is that it is likely only episodes of imagining from the inside that involve changing one’s point of view that involve this phenomenology of ‘distancing’. This is certainly the case in the example: the subject imagines observing a beautiful scene; within that episode of imagination, the subject imagines taking up a different point of view, at a later point in time, talking to a friend. In this case, the imagined act of talking to a friend would seem ‘distant’, at least phenomenologically. To state the obvious, though, if it were possible to imagine from the inside imagining speaking from the inside (without a change of point of view), such an episode would not involve a change in point of view. Perhaps, then, if

---

<sup>37</sup> Thanks to an anonymous reviewer at the *Review of Philosophy and Psychology* for this suggestion.

<sup>38</sup> Thanks to David Chalmers for this suggestion.

one thinks that cases of the latter kind do involve the phenomenology of distancing, then one has just made an introspective error: one was, in fact, somehow imagining from the inside imagining speaking from the inside *with* a change in point of view.

In brief, then, neither the phenomenon of collapsed iteration, nor the phenomenology of distancing, can usefully be appealed to in challenging the notion that it is not possible to imagine from the inside imagining performing some action from the inside.

### 3.2. The Parallel Cases Argument

The Parallel Cases Argument demonstrates that inner speech stands in the same relation to imagined speech as external speech stands in relation to speech produced in a pretend scenario. It contends that this provides a reason to treat external speech and inner speech alike, i.e. as kinds of actual speech.

#### 3.2.1. Two Cases

Consider the following two simple cases:

Case 1: An actor in a play is delivering his lines. He is playing Polonius in a production of *Hamlet* and saying ‘This above all: to thine own self be true’. Suddenly he stops, abandons character and makes an observation about the play. He says, ‘I never know what to make of this character. He’s a self-important old chap but he has some wonderful lines.’ Then, again suddenly, he resumes performing his part, ‘And it must follow, as the night the day ...’

Case 2: A politician is imagining giving a speech. She is rehearsing a campaign speech and imagining saying ‘Healthcare in this country has been neglected for too long’. Suddenly, she pauses and makes an observation to herself in inner speech about the speech she is imagining giving. She utters in inner speech, ‘I have to remember to address the audience directly’. Then, again suddenly, she resumes imagining giving her speech, ‘You and your children are entitled to quality care ...’

How should we describe Case 1? Initially, the actor is producing speech within a kind of make-believe scenario. Let us call this kind of speech ‘pretend speech’. He then produces an interpolation, which is not part of the make-believe scenario, even if (improbably) this were not immediately apparent to the audience. It is, we are surely inclined to say, an utterance in



actual speech, in contrast to the pretend speech which is a part of the play. Finally, he resumes producing pretend speech.

But what about Case 2? The politician begins by producing imagined speech. One way we can think of imagination is as an internal version of pretence.<sup>39</sup> An imagined action takes place in the mind while a pretend action takes place in the external world, but each resembles an action that could actually be performed. On the face of it, the politician's interpolation stands in the same relation to her imagined speech as the actor's interpolation does to his pretend speech. The actor and the politician both produce comments on what they are pretending or imagining saying. Accordingly, insofar as we want to treat the politician's interpolation in external speech as an instance of actual speech, it seems we must also treat the actor's interpolation in inner speech as an instance of actual speech, albeit a silent one. At the end, the politician resumes producing imagined speech.

The parallel between the contrasts in Case 1 and Case 2 can be emphasised by recalling a point that was made in developing the Actual Speech View. There, it was urged that, if we had direct access to one another's inner speech (i.e. telepathic access, as it were), then inner speech would facilitate communication and enable performative utterances, just as external speech does. Now, suppose we had such direct access to the utterances the politician produces. It seems clear that we would want to draw the same distinction between her imagined speech and her inner speech as we do between the pretend speech and the external speech of the actor. It would make no difference at all to us that her utterances were produced silently, while the actor's utterances were produced aloud.

### **3.2.2. Possible Speech**

The Imagined Speech View, as it was set out, treats inner speech as possible speech. Producing inner speech involves imagining saying something that one could, at least in principle, say externally, i.e. in actual speech. Noting this, one might object to the analysis of Case 2 on the following terms. Just as the initial utterance that the politician imagines producing is an utterance that she could produce in external speech, so the interpolation that she produces in inner speech is an utterance that she could produce in external speech. Accordingly, one might think that the interpolation produced in inner speech, as an utterance that one could produce externally, is in fact an utterance in imagined speech.

---

<sup>39</sup> Thanks to Daniel Stoljar for suggesting this.

However, the Actual Speech View can also allow that inner speech is speech that one could produce externally. There is no problem in supposing that an utterance produced in actual speech which is silent can also be produced in actual speech which is audible. Accordingly, the fact that the politician's interpolation in inner speech in Case 2 could also be produced aloud does not interfere with the conclusion that it is an instance of actual speech.

### **3.3. The Continuity Argument**

The Continuity Argument shows how inner speech can be seen as sitting on a continuum with various kinds of external, and therefore actual, speech. It urges that this is best explained by the theory that inner speech is also a kind of actual speech.

#### **3.3.1. Continuous Phenomena**

Consider the following different types of linguistic activity. First, there is inner speech, which involves auditory imagery of speech and possibly some minimal movements in the muscles which are involved in producing external speech.<sup>40</sup> Secondly, there is soft, semi-vocalised mumbling or speaking under one's breath, without any interlocutor – 'to oneself', as it were. Thirdly, there is fully vocalised speech to oneself. These latter two are sometimes referred to as 'private speech'. When the different types of linguistic activity are arranged this way, it is very tempting to see them as continuous phenomena: phenomena of the same kind, varying along a couple of dimensions such as volume and the amount of physical movement which is involved. And, insofar as we would think of semi-vocalised and fully vocalised private speech as kinds of actual speech, we would surely incline to think that inner speech is also a kind of actual speech.<sup>41</sup>

Still, there is an alternative explanation: that we are imagining producing private speech when we produce inner speech. So why prefer the hypothesis that inner speech is a kind of actual speech, continuous with external, private speech?

The following reason suggests itself. It is a familiar, albeit occasional experience to slip inadvertently from producing inner speech to producing audible private speech, most likely of the soft and semi-vocalised variety. If inner speech were imagined speech, then this is not something that we would expect to happen. We do not ordinarily slip inadvertently from merely imagining doing something to actually doing it. It seems more comfortably

---

<sup>40</sup> Though, again, this is not clear: see Section 2 in the Introduction to the thesis.

<sup>41</sup> Thanks to Daniel Stoljar for suggesting this argument to me.

accommodated on the thesis that inner speech is a kind of actual speech, differing from the external versions of private speech only in volume and the amount of physical movement involved. On this thesis, slipping from producing inner speech to producing soft, semi-vocalised private speech is no more mysterious than slipping from producing soft, semi-vocalised private speech to producing fully vocalised private speech.

### **3.3.2. Imagination and Performance**

A possible objection to the Continuity Argument might run along the following lines. It is actually not so unusual to slip inadvertently from imagining performing an action to actually performing it. Consider this case, which seems plausible. I am sitting across a table from some interlocutor who is infuriating me. Increasingly enraged, I start to imagine throwing a glass of water in his or her face. Eventually, I lose control and actually do throw the glass of water. Is this not a straightforward case of slipping inadvertently from imagination to action?<sup>42</sup> Another example would be a sportsperson – perhaps a gymnast – imagining a part of their routine before trying to perform it.

The phenomenon which these cases exemplify is familiar, but it is a different phenomenon. It involves imagining performing an action before actually performing the same action. This is not the same as producing speech internally and then continuing one's monologue – producing different utterances – externally. The latter phenomenon still seems best explained on the hypothesis that there is a continuum of kinds of actual, private speech, including inner speech, and that one can move up and down this continuum, as it were.

### **3.4. The Precisification Argument**

The Imagined Speech View provides a fairly detailed account of the type of imagined speech that inner speech might be. To recap: on the Imagined Speech View, inner speech involves imagining oneself speaking rather than imagining another person or other people speaking. It involves imagining performing (rather than observing) the act of speaking. To say the same thing with a different emphasis, it involves imagining speaking from the inside rather than from the outside. Producing inner speech involves producing auditory imagery and it may involve some activity in the muscles associated with external speech production. But this leaves a lot unspecified. For example, does inner speech involve imagining speaking to someone other than oneself, whether a particular person or some generic interlocutor? Does it

---

<sup>42</sup> Thanks to Jennifer Corns for suggesting both the objection and the example here.

involve imagining speaking in a particular way, e.g. such that only you can hear what you say? Does it involve imagining saying something in a particular context?

The Precisification Argument urges that there is no viable way in which the Imagined Speech View can be precisified so that questions such as those above can be answered. It considers a range of possibilities which seem to have some promise and argues that none of them is adequate. It acknowledges that not every possible precisification of the Imagined Speech View can be considered in advance. By contrast, though, it emphasises that there is no equivalent problem of precisification for the Actual Speech View.

So, how might the Imagined Speech View be made more precise? The following answers seem like they have the most potential:

1. Producing inner speech involves imagining saying something in precisely the situation that you are now in;
2. Producing inner speech involves imagining saying something in such a way that only you can hear it;
3. Producing inner speech involves imagining saying something to a particular interlocutor;
4. Producing inner speech involves imagining saying something which is not directed to anyone;
5. Producing inner speech involves imagining saying something to yourself; and
6. Producing inner speech just involves imagining saying something, without any imagined context.

At this point, one might worry that the Precisification Argument depends on an unjustified assumption. Supposing inner speech is a kind of imagined speech, why assume that it is a kind of imagined speech which can be described with any degree of precision? Could it not be that the kinds of features which are presupposed by questions such as those above – e.g. does inner speech involve imagining speaking to someone other than oneself? – are simply indeterminate in the relevant kind of imagined speech? It is notorious, after all, that not everything we imagine is determinate. Daniel Dennett (1969) demonstrates this fact with the example of ‘the Tiger and his Stripes’ (p. 136). He asks: ‘I can ... imagine ... a striped tiger, but must the tiger I experience have a particular number of stripes?’ (1969, p. 136). The answer, we feel compelled to say, is ‘no’. Even if you try to count the stripes on an imagined tiger, it is most unlikely that you will confidently arrive at a particular number. Still, there is no doubt that it is a tiger that has been imagined. Might one not think, then, that producing inner speech involves imagining speaking in a way which leaves everything else indeterminate? To address this concern, another suggestion should be added to the list:

7. Producing inner speech involves imagining saying something in a way which leaves everything else indeterminate.

The issue of indeterminacy in imagination, however, will also have relevance in considering Suggestions 3, 4 and 5.

Let us now work through these possibilities, and see why none is satisfactory. When we come to them, Suggestions 4 and 5 will be treated together.

#### **3.4.1. Suggestion 1**

Suggestion 1 (producing inner speech involves imagining saying something in precisely the situation that you are now in) is a natural first suggestion. It may also seem quite appealing, just because it is minimal and simplistic. It does not obviously entail that inner speech is more complicated than its familiar phenomenology suggests – which, as we will see, is a problem for some of the more sophisticated suggestions that follow.

However, there is a major problem with Suggestion 1 which is relatively easy to see. Take a case in which the situation you are now in involves the presence of other people. If producing inner speech involves imagining saying something in precisely the situation that you are now in, then, in such a case, producing inner speech would involve imagining saying something in such a way that those others present could hear it. But this is not a feature of inner speech. Consider the case in which, during a conversation with a difficult brother-in-law, you produce the inner speech utterance, ‘This guy is such a jerk’. In producing this inner speech utterance, you do not imagine that you brother-in-law can hear it.

#### **3.4.2. Suggestion 2**

Suggestion 2 (producing inner speech involves imagining saying something in such a way that only you can hear it) also has some intuitive appeal, perhaps especially if it is combined with Suggestion 1 (producing inner speech involves imagining saying something in precisely the situation that you are now in). One might think that inner speech involves imagining saying something in precisely the situation that you are now in, but imagining saying it in such a way that only you can hear it. As noted just above, Suggestion 1 by itself is subject to the objection that imagining saying something in any situation in which others are present would imply imagining that those others could hear what you say. Suggestion 2 and Suggestion 1 taken together avoid this problem.

However, a different problem for Suggestion 2 (even if combined with Suggestion 1) becomes apparent when we consider more closely exactly what it would involve. There are two possibilities. First, one might think that imagining saying something in such a way that only you can hear it amounts to imagining whispering or speaking extremely softly. But, surely, we do not ordinarily feel as if we are doing anything like whispering when we produce inner speech. We do not ordinarily experience our inner speech as having some kind of hushed tone, nor do we have a sensation that it is only just audible even to ourselves.

Alternatively, one might think that imagining saying something in such a way that only you can hear it amounts to imagining speaking in some special way such that others cannot hear, no matter how close they might be physically and how acute their hearing – i.e. in some way that it is not possible even in principle for others to hear. As soon as it is put like this, though, we can see that it is not really a plausible characterisation of inner speech. It would involve imagining speaking in some special way in which we are not actually able to speak externally, and this also does not seem to describe any aspect of the phenomenology of inner speech. (Actually, it is not at all clear that we even can imagine speaking in this special way. Suppose you were asked to do so. What exactly would you try to imagine?)

### **3.4.3. Suggestion 3**

Suggestion 3 is that producing inner speech involves imagining saying something to a particular interlocutor. If this were right, then we would expect the following to hold. After producing some utterance in inner speech, you should be able, if asked, to provide some information about the distinct individual to whom you imagined you were speaking. Who was the imagined interlocutor? Can you perhaps describe this individual you imagined speaking to? Or nominate at least one attribute you imagined them having? But this is surely not the case. You would not be able to answer any such questions after producing an utterance in inner speech.

Here, borrowing the point which (I foreshadowed) motivates Suggestion 7, it might be urged that inner speech does involve saying something to an imagined interlocutor but that all of the features of this imagined interlocutor are determinate. When we produce inner speech, we do imagine speaking to some interlocutor but we are not able to provide further detail about this interlocutor, just as we are not able to nominate the number of stripes on an imagined tiger.

The problem with the suggestion that details about the imagined interlocutor might be indeterminate, however, is that it is not just some details which are indeterminate. After producing an utterance in inner speech, one would not be able to answer any questions of the

kind above. And if one cannot point to even a single attribute that they take an imagined interlocutor to have, then surely this can only be because they did not imagine any interlocutor. One cannot imagine something which has no features at all. Compare: one cannot imagine a tiger and be unable to indicate not only the number of stripes it has but also its colour, its shape, and everything else.

It is worth noting that this response could also be given if it were suggested that producing inner speech involves imagining speaking not to a particular interlocutor but to some generic, largely indistinct interlocutor. It is still the case that, after producing an utterance in inner speech, one would not be able to nominate any attributes that they took the imagined interlocutor to have. And, if inner speech did involve imagining speaking to some generic, largely indistinct interlocutor, then we would still expect that, after producing an utterance in inner speech, it would be possible to nominate at least some attribute one took that interlocutor to have. Again, compare: one cannot imagine even a generic, indistinct tiger and be unable to indicate not only the number of stripes it has but also its colour, its shape, and *everything* else.

#### **3.4.4. Suggestions 4 and 5**

The first thing to note about Suggestion 4 (producing inner speech involves imagining saying something which is not directed to anyone) and Suggestion 5 (producing inner speech involves imagining saying something to yourself) is that they are largely the same suggestion, put differently. It is common sometimes to produce utterances in external speech without any distinct listener, e.g. if you think you are alone or that no one else is paying attention, perhaps if you are deep in thought while working through some puzzle. And it seems natural to describe such an event either as saying something which is not directed to anyone or as saying something to oneself. A parallel point applies if you imagine producing such an utterance: it would be natural to describe the episode either as imagining saying something which is not directed to anyone or as imagining saying something to yourself. (There is a small range of cases which provide a reason to believe the two suggestions can be separated; we will come to these shortly.)

Now, one certainly can imagine producing utterances of this kind in external speech. Is this what inner speech is? As was shown in considering Suggestion 3, inner speech does not involve imagining speaking to an interlocutor – either a particular interlocutor or some generic, largely indistinct interlocutor. So, if inner speech is imagined speech, then presumably it does involve imagining producing the type of utterances which we would describe as being directed either to oneself or to no one in particular.

But we quickly come to a problem. Again, it is common to produce a kind of external speech which is naturally described either as saying something which is not directed to anyone or as saying something to oneself. This is private speech. So, to imagine producing utterances of this kind is to imagine engaging in private speech. But private speech takes place in a context. One is always somewhere, in some situation, when one produces private speech. So what context does one imagine being in, if producing inner speech amounts to imagining producing private speech?

The answer which suggests itself most readily is that one just imagines producing private speech in the context one is now in. But this is very similar to Suggestion 1: producing inner speech involves imagining saying something in precisely the situation that you are now in. And the problem which arose for Suggestion 1 also arises here. If one's current situation is one in which other people are present, then imagining producing private speech in one's current situation would involve imagining saying something such that those other people could hear it – and this is clearly not a part of inner speech.<sup>43</sup>

Another possibility is that producing inner speech involves imagining producing private speech in some context other than the context one is now in. Certainly, it is possible to produce inner speech while, for example, imagining seeing something which is not presently within one's sight.<sup>44</sup> Take this case. A young equestrian is thinking about a horse she is considering buying. She visualises the animal from the side and produces the inner speech utterance, 'He's so beautiful'. But this does not amount to imagining producing private speech in a particular context. That would require that the young equestrian imagines, for example, producing the utterance aloud while standing near the animal in a paddock. But episodes such as this – where one actually imagines *being* in some other context – seem far more elaborate than what we ordinarily think of as inner speech.<sup>45</sup>

One last possibility is that producing inner speech involves imagining producing private speech but with the details of the context left indeterminate. But we already know what problem will arise here as well. Suppose you produce an utterance in inner speech. If asked,

---

<sup>43</sup> Of course, it is true that we often, perhaps typically, produce private speech when we are alone and no one else can hear what we say. And, when actually in such a context, imagining producing private speech in one's present context would not involve imagining that anyone else can hear what one says. Still, we can and sometimes do produce inner speech when we are in the company of others. And imagining producing private speech when we are in such a context would involve imagining that others can hear it.

<sup>44</sup> There will be significantly more about this in Chapter Two.

<sup>45</sup> Of course, one can imagine producing *inner* speech in some context other than one's present context. We saw a couple of examples of this earlier. One is the case from Section 3.1.3, involving the golfer who imagines being on a putting green and, within that context, producing inner speech. However, as discussed there, this is a case of imagining producing inner speech, not a case of producing inner speech.



you would surely not be able to provide any details about the supposed imagined context, which rules out the possibility that there is an imagined context at all.

It was foreshadowed parenthetically that there is a small range of cases which provide a reason to believe that Suggestion 4 (producing inner speech involves imagining saying something which is not directed to anyone) and Suggestion 5 (producing inner speech involves imagining saying something to yourself) can be separated. These are cases in which one issues an imperative to oneself in inner speech. Recall the anecdote regarding Jerry Fodor from the Introduction to the thesis. He allegedly produces inner speech utterances such as ‘C’mon, Jerry, you can do it’ when thinking particularly hard. Or take a case in which a tennis player, exasperated with her disappointing performance, produces an inner speech utterance such as ‘Pay attention!’ These are cases where it is plausible that the individual is imagining saying something to him/herself but implausible that they are imagining saying something which is not directed to anyone.

The difficulty with this point is that imperatives issued to oneself in inner speech actually seem like cases which are especially troublesome for the Imagined Speech View. Merely imagining saying ‘Pay attention!’ seems like a particularly fruitless thing for a frustrated tennis player to do. The tennis player wants to motivate herself, or influence her own behaviour in some way, just as she might do by yelling the imperative aloud. And she would scarcely motivate herself by imagining doing this. This is easily tested: just imagine issuing an imperative to yourself next time you find yourself annoyed with your performance in any activity. The difference between doing this, and issuing the imperative in inner speech, will, I suggest, be palpable.<sup>46</sup>

#### **3.4.5. Suggestion 6**

Suggestion 6 is that producing inner speech just involves imagining saying something, without any imagined context (i.e. with no imagined context at all, not even a context with indeterminate features). We do sometimes imagine things in this way. One might, for example, just imagine a jar of peanut butter without any background. Still, there is a problem with Suggestion 6, though it is a little difficult to see.

It is true that, if you try to imagine saying something without any context, you will likely just find yourself producing inner speech. Suppose, for example, that you try to imagine saying ‘Grass is green’ without any context. It would be surprising if you did not just produce the

---

<sup>46</sup> Thanks to Frank Jackson for the points in the last couple of paragraphs, as well as the tennis example.

utterance, 'Grass is green', in inner speech. But the reason for this might not be that this kind of imagined speech is what inner speech amounts to. The reason may simply be that you tried to produce the relevant kind of imagined speech but did not succeed, i.e. that you tried to produce the relevant kind of imagined speech and instead produced inner speech.

Given this difficulty, how can we see that producing inner speech does not just amount to imagining saying something without any context? I suggest that we can imagine speaking in the relevant way, and see how this differs from producing inner speech, by first imagining speaking in a much more elaborate way, and then gradually paring back the additional features within that imagined episode.

Let us start with a familiar example. Imagine performing the act of giving the State of the Union Address before a House of Representatives full of dignitaries, with your voice amplified and reverberating around the chamber. Now imagine giving the address before a House of Representatives full of dignitaries, with your voice amplified but not reverberating around the chamber. Now imagine giving the address before a House of Representatives full of dignitaries, but without your voice amplified or reverberating around the chamber. And now imagine just giving the address, without the audience. In this final case, one is just imagining speaking without any context. And there is still, I suggest, a phenomenological difference between doing this and producing inner speech.

This phenomenological difference is, admittedly, slight. Possibly, it would become clearer if one practised the process of coming to imagine speaking in this way. Still, it does seem like there is some such difference, which is enough to separate the two phenomena.

#### **3.4.6. Suggestion 7**

Suggestion 7 is that producing inner speech involves imagining saying something in a way which leaves everything else indeterminate. The idea is that producing inner speech involves imagining speaking but not in any way that we are able to describe in detail. Accordingly, it is mistaken to expect that we would be able to answer specific questions about, for example, whether producing inner speech involves imagining speaking to an interlocutor, or in a particular context, or even in no context at all.

We can see why this suggestion will not work by recalling again the example of the tiger and its stripes. When you imagine a tiger, it is indeterminate how many stripes the tiger has: you can say that it has stripes, but you cannot say how many. But note that, while you cannot confidently say how many stripes the tiger does have, nor can you confidently say that it does not have a particular number of stripes. You cannot say that it has 16 stripes, but nor can you

say that it does not; you cannot say that it has 17 stripes, but nor can you say that it does not; and so on.

It is different with inner speech. It is not that we cannot say whether we are imagining speaking to an interlocutor or imagining producing private speech (i.e. speaking without an interlocutor) when we produce inner speech; we can say that neither of these things is the case. It is not that we cannot say whether we are imagining speaking in our current situation or in another context or in no context at all when we produce inner speech; we can say that none of these things is the case. We have seen above the reasons for each of these points.

So it cannot be that producing inner speech involves imagining speaking in a way that leaves everything else indeterminate. If producing inner speech does involve imagining speaking, then there are some questions which immediately suggest themselves, such as whether it involves imagining speaking to someone, and whether it involves imagining speaking in a certain context. We have seen, though, that it is not because of indeterminacy that questions such as these cannot be answered.

#### **3.4.7. Completing the Argument**

A proponent of the Imagined Speech View might protest that we have not considered every possible way that their view might be precisified. This is so, and it may be that some compelling suggestion, which circumvents the problems besetting the suggestions above, could be made. Still, several apparently promising proposals have now been considered, and none has survived very much scrutiny at all.

By contrast, there is simply no problem of precisification for the Actual Speech View. On the Actual Speech View, producing inner speech involves producing silent, private speech in the situation that you are now in. However, only you can hear your inner speech, as it were, just because it is silent. Accordingly, there is no problem of imagining that anyone else who happens to be in your vicinity might hear what you say in inner speech. Inner speech does not involve imagining speaking aloud but in some way such that it is not audible to others; it just involves speaking in a way that is not audible at all. Producing inner speech does not involve speaking to an interlocutor, any more than producing audible private speech involves speaking to an interlocutor (although it is true that others may hear one's audible private speech). There is a context in which it takes place. As already noted, producing inner speech just amounts to speaking silently in your present context.

The contrast is instructive. The problem of precisification represents a major challenge for the Imagined Speech View. There is no parallel problem for the Actual Speech View.

## **4. Concrete Tokens**

I foreshadowed at the end of Section 2 that I would consider in Section 4 a point that might be made either by a proponent of the Imagined Speech View in arguing against the Actual Speech View, or by a proponent of the Quasi-Speech View in an effort to separate their view from the Actual Speech View. In broad terms, this point is that producing external speech, but not inner speech, essentially involves producing some kind of concrete, physical token. Accordingly, insofar as we can think of external speech as the paradigm of actual speech, this might be a reason to believe that inner speech is not a kind of actual speech after all.

### **4.1. The Concern**

As we saw in setting out the competing characterisations of inner speech, a proponent of the Imagined Speech View can emphasise that producing external speech involves vocalising an utterance whereas producing inner speech does not. When one produces external speech, one actually generates a sound which can be heard by anyone listening. One might think that this feature is essential to actual speech. After all, it is what allows external speech to perform some of the functions which we most readily associate with it, such as facilitating communication and enabling performative utterances. And it is a feature which is missing from inner speech.

A proponent of the Actual Speech View is not defenceless against this claim. As we saw, it might be urged in response that it is only contingently true that speech needs to be vocalised in order to serve the kinds of functions that we most readily associate with it. If we did somehow have unmediated access to one another's inner speech – and, again, there is nothing a priori inconceivable about this kind of telepathy – then inner speech would be facilitative of communication and it would enable performative utterances.

But the point that external speech involves vocalising an utterance can be developed further by recruiting the type-token distinction. On the standard view, the production of a linguistic item always involves the generation of some concrete, physical token which instantiates a linguistic type. The difference is often explicated in the following way. Take the following series of words: 'knife, fork, spoon, knife, fork, spoon'. How many words are there in that series? The question can be answered in two ways. There are three word types: 'knife', 'fork', and 'spoon'. But there are six word tokens, as each word type is instantiated twice. Word tokens are concrete instances and they can take a wide range of forms: sounds or hand

movements, some inscription in print, or, in the case of Braille, a series of cells, and so on. Critically, both a word token, and a word type which it instantiates, are necessary for a word to be spoken, written etc. Obviously enough, there must be some physical entity – sounds, ink printed on paper etc. But this physical entity must also instantiate a word type. A jumble of letters typed on a page – ‘Idienfo’ – is not a word, precisely because it does not instantiate a word type. All of this holds, *mutatis mutandis*, for sentence types and sentence tokens, paragraph types and paragraph tokens, and so on.<sup>47</sup>

It might be argued that it is because producing inner speech does not involve producing concrete, physical entities instantiating linguistic types that inner speech is not a kind of actual speech. No sound, no print is generated when one produces inner speech. Set aside the question whether producing such physical entities is necessary for speech to facilitate communication or enable performative utterances or whatever else. What is really significant, the argument goes, is that producing concrete tokens which instantiate linguistic types is necessary for actually generating linguistic items, and generating linguistic items, in turn, is surely necessary for actually speaking. This is why it matters that inner speech is not vocalised: if it were, then producing it would involve producing physical entities.<sup>48</sup>

#### **4.2. Two Difficulties**

As foreshadowed, the issue just discussed could be recruited either by a proponent of the Imagined Speech View in arguing against the Actual Speech View, or by a proponent of the Quasi-Speech View seeking to separate their view from the Actual Speech View. It is straightforward to see the role it would play in both of these strategies.

A proponent of the Imagined Speech View would simply urge that inner speech cannot be actual speech because producing inner speech does not involve producing actual physical entities. By contrast, the Imagined Speech View can easily accommodate the fact that producing actual speech involves producing physical entities. On the Imagined Speech View, producing inner speech involves imagining speaking aloud – and therefore imagining producing physical entities.

A proponent of the Quasi-Speech View would claim that a critical feature of actual speech which inner speech lacks has now been found and that it should therefore be concluded that, while inner speech has many of the features of actual speech, it does not have them all. For it is not open to a proponent of the Actual Speech View to downplay the fact that inner speech

---

<sup>47</sup> We find the terminology of ‘type’ and ‘token’ as early as Peirce (Peirce 1933, p. 423).

<sup>48</sup> Thanks to Daniel Stoljar for pressing me on this point.

lacks this feature. It is not open to say that the difference would disappear if we had unmediated access to one another's inner speech. This, unlike some of the other features of external speech which inner speech lacks – e.g. that it facilitates communication and enables performative utterances – would persist even if we did have unmediated access to one another's inner speech.

So there is a challenge that an advocate of the Actual Speech View really needs to meet, on pain of making a damaging concession to both an advocate of the Imagined Speech View and an advocate of the Quasi-Speech View.

### **4.3. Meeting the Challenge**

Recall why one might think it is important that producing inner speech does not involve producing concrete physical tokens. The issue is that producing concrete tokens which instantiate linguistic types is necessary for actually generating linguistic items, and generating linguistic items, in turn, is surely necessary for actually speaking. I suggest that there are three points that a proponent of the Actual Speech View can make in response.

#### **4.3.1. Non-Physical Tokens**

David Kaplan (1990) casts doubt on whether tokens of word types must be physical entities. He makes the point that a word can be stencilled (1990, p. 97). That is, the letters of a word may be cut out of a piece of paper. In this case, the word will be tokened not by the piece of paper or anything inscribed on it, but by the vacant space removed from the page. Kaplan explains that this point occurred to him when thinking about words inscribed into stone. The token is not the stone; it is the space removed from it (1990, p. 97). And there are other examples of this kind, such as a name carved into bark or a message written in wet sand. *Perhaps* it could be claimed that there is still a physical object in these cases; the physical object is the vacant space. But this is very dubious, especially if we are just dealing with the everyday sense of physical object, i.e. something which takes up space. In these cases, it seems that word types are tokened by *empty* space.

### 4.3.2. Intuitive Judgements

Think about some specific cases:

1. You are reading a newspaper article aloud to your elderly mother whose eyesight has deteriorated to the point that she is no longer able to read it for herself. Noticing that she has fallen asleep, you continue reading the article for your own interest, but silently, in inner speech. Have you stopped actually reading the words?
2. A poet is composing a short poem in inner speech. Before she has committed it to paper, does the poem already exist? (Wetzel 2014, sec. 6).
3. A primary school student is looking over the maths homework she must do. She does a few of the sums in her head, so to speak. Has she not been doing arithmetic?

What shall we say about these cases? Regarding Case 1, we surely think of reading aloud as involving the generation of tokens of linguistic types, guided by the printed tokens on a page. But if this is what you do when you read aloud, it is at least plausible that you continue doing just that when you switch from reading aloud to reading silently. In any event, if one had an initial inclination to think that producing inner speech involves imagining speaking and merely imagining producing physical tokens of linguistic types, that inclination may very well be weaker in the case of silent reading.

In Cases 2 and 3, the intuition that linguistic items are being tokened is probably stronger. The intuitive analysis of Case 2 is surely that the poem does exist, even if it is not yet recorded in some external medium; the poet has not merely imagined it (Wetzel 2014, sec. 6).<sup>49</sup> Yet a poem is a linguistic construction; it consists of linguistic types instantiated by linguistic tokens. Regarding Case 3, it would be extremely surprising if one thought that the student is not doing arithmetic. Doing arithmetic involves tokening a kind of word, i.e. the kind of word we use for numbers. And the student who we take to be tokening these word items is doing so mentally; she is not writing anything down.<sup>50</sup>

The point of all of this is to show that the received view that linguistic items involve linguistic types instantiated by concrete, physical tokens is not entirely refined. At least, it does not entirely match our pre-theoretical judgements about language.

---

<sup>49</sup> Wetzel writes that there is 'disagreement' over this analysis but does not provide references (Wetzel 2014, sec. 6).

<sup>50</sup> Thanks to Frank Jackson for this example and analysis of it.

### **4.3.3. Linguistic Images**

Consider a photograph of a sentence written on a piece of paper. There is a photographic image of a sentence token, i.e. an image of the sentence written on paper. But there is not only an image of a sentence token; there is also an actual sentence token, i.e. the photographic image itself. One can read it, spot spelling mistakes etc. But the actual token of the sentence type does not exist separately from the photographic image of the sentence token written on paper. It has the same physical realisation. Similarly, consider a recording of someone saying a word. As well as the auditory image of the word, the word type is actually being tokened. Again, though, the token of the word type does not exist separately from the sound waves which are generated as the recording is played.

Once we notice that this holds for external images of linguistic items, it is hard to see why we should not also say that mental images can be linguistic tokens as well. When one has a visual image of a sentence, the visual image itself is a token of the sentence type; when one utters a sentence in inner speech and produces auditory imagery of that sentence, the auditory image is itself a token of the sentence type.<sup>51</sup>

### **4.3.4. Combining the Responses**

There is reason to doubt that tokens of linguistic types have to be concrete, physical entities; our pre-theoretical judgements allow that linguistic types can be tokened in the mind; and we have an understanding as to how linguistic types might be tokened when one produces the auditory imagery involved in inner speech. This should allay the concern that inner speech does not involve the generation of concrete, physical tokens entirely. It is not open to a proponent of the Quasi-Speech View to claim that an essential feature of actual speech which is missing from inner speech has been found, and therefore to separate their view from the Actual Speech View. And a proponent of the Imagined Speech View gains no advantage in pointing out that it is possible to imagine producing concrete, physical tokens, because tokens of linguistic types need not be concrete, physical entities.

## **5. Assessing the Case**

We have now seen four arguments which support the view that inner speech is a type of actual speech and set aside an objection to the view. The Embedding Argument emphasises that we

---

<sup>51</sup> Thanks to Daniel Stoljar for suggestions here.



engage in inner speech – like external speech but unlike speech imagined from the inside – without adopting a new point of view. The Parallel Cases Argument shows how inner speech appears to stand in the same relation to imagined speech as actual external speech stands in relation to what I called ‘pretend speech’, and suggests that inner speech and actual external speech should therefore be treated alike. The Continuity Argument suggests that inner speech is best seen as standing on a spectrum with both semi- and fully-vocalised private speech, which are clearly kinds of actual speech. And the Precisation Argument reviews a wide range of suggestions as to the particular kind of imagined speech that one might think inner speech consists in, but shows that none of these suggestions succeeds. How far do these arguments take us?

### **5.1. The Two Views**

Recall once more the competing ways that inner speech is characterised by the Imagined Speech View and the Actual Speech View. On the Imagined Speech View, producing inner speech involves imagining performing the action of speaking from the inside, in a way which involves a relatively bare sensory phenomenology. On the Actual Speech View, inner speech is private, actual speech, which we produce silently. The first thing to note is that the arguments have force against the Imagined Speech View, or for the Actual Speech View, as those views were originally explicated.

The Embedding Argument is an argument against the Imagined Speech View. It turns on the point that, on the terms of the Imagined Speech View, producing inner speech involves imagining speaking from the inside. The key observation is that it is possible for an episode of inner speech to be embedded within an episode of imagining from the inside, but that it is not possible for an episode of imagining speaking from the inside to be embedded within another episode of imagining from the inside. This implies that engaging in inner speech is not the same thing as imagining speaking from the inside.

The Parallel Cases Argument and the Continuity Argument support the Actual Speech View directly. The Parallel Cases Argument shows that inner speech and external private speech stand in the same relationship to imagined and pretend speech, respectively. It concludes that inner speech should therefore be treated as a kind of actual speech, just as external private speech is. The Continuity Argument shows how inner and external private speech are most naturally seen as graded phenomena, differing only in volume and the amount of physical movement involved in producing them.

The Precisification Argument is primarily an argument against the Imagined Speech View but it also provides some direct support for the Actual Speech View. Key to this argument is that, while the Imagined Speech View provides a reasonably detailed characterisation of inner speech, it leaves a lot unspecified. If we try to make the view more precise so that a proponent of the view can answer questions which it seems reasonable to ask, serious difficulties emerge. By contrast, there is simply no problem of precisification for the Actual Speech View.

## **5.2. Reviving Imagined Speech?**

Before setting out the Imagined Speech View, I indicated that I would try to present the best possible version of the view that inner speech is imagined speech, but that it may be that one could disagree, at least with respect to some of the details. We have just seen that the Parallel Cases Argument and the Continuity Argument offer direct support for the Actual Speech View. However, the Embedding Argument is only an argument against the Imagined Speech View and the Precisification Argument is primarily an argument against it. This raises the question: could the view that inner speech is imagined speech be developed in some way such that those arguments would not have force against it? To answer this, we first need to see what it is that makes the Imagined Speech View vulnerable to the Embedding Argument and the Precisification Argument.

As noted in the last subsection, the Embedding Argument turns on the point that, if inner speech involves imagining speaking, then it involves imagining speaking from the inside. The Precisification Argument turns on the point that, if inner speech involves imagining speaking, then it involves imagining performing the action of speaking. The suggestions that were considered were suggestions as to whether producing inner speech might involve imagining speaking in a particular context; or imagining speaking in such a way that only you can hear yourself; or imagining speaking in a particular situation; or imagining speaking in a way that leaves everything indeterminate (except that one is imagining speaking). All of these are just more detailed suggestions about what might be involved when one imagines performing the action of speaking.

Now, as we saw early in this chapter, the claim that producing inner speech involves imagining speaking from the inside, and the claim that it involves imagining performing the action of speaking, are really just different ways of expressing the same idea. And it really is very difficult to see how the idea that inner speech is a type of imagined speech could be developed without this idea. The suggestion that inner speech might involve observing speech from the outside is implausible. Accordingly, it is hard to see how anyone who believes that inner

speech is a kind of imagined speech could avoid the Embedding Argument and the Precisification Argument.

Is there anything at all left for someone who believes that inner speech is a kind of imagined speech? There is one possibility. One might take the view that any mental state involving mental imagery – involving some conscious perceptual representation – is an instance of imagination, and it is clear that inner speech involves conscious perceptual (auditory) representations.<sup>52</sup>

The position is certainly coherent: there is nothing illogical about it. However, there are a couple of points that should be made about it. First, it is drastically weaker than any version of the theory that inner speech is a type of imagined speech which might initially have seemed attractive – e.g. something like the Imagined Speech View. Indeed, in setting out the Imagined Speech View, it was noted that inner speech does involve imagery (of a relatively bare kind), but that this does not seem like anything approaching a comprehensive characterisation of the phenomenon.

Secondly, the observation that inner speech involves mental imagery is entirely compatible with the Actual Speech View. In fact, it is an explicit commitment of that view. A proponent of the Actual Speech View believes that inner speech is a kind of actual speech notwithstanding that it involves the production of imagery rather than, say, sounds. So, if one can only insist that inner speech is imagined speech because it involves mental imagery, then one is making a weak claim indeed. Really, if someone who thinks that inner speech is imagined speech retreats this far, they have given up on the dispute.

## **6. Some Implications**

The problem which has been the focus of this chapter should be of interest to anyone who wants to understand inner speech, for it goes directly to the nature of the phenomenon. In closing this chapter, however, I want to mention three other respects in which the status of inner speech as actual speech is both important and interesting.

### **6.1. Internalisation**

At the end of the Introduction to the thesis, I summarised Vygotsky's theory on the development of inner speech. He believed that inner speech develops as egocentric speech –

---

<sup>52</sup> Thanks to Alma Barner for this suggestion.

a kind of private speech that children produce which seems to be directed towards whatever activities are occupying their attention – is internalised, or ‘goes underground’, as he puts it. I also noted that it is not clear what this ‘internalisation’ might involve. How exactly does an activity which one performs externally become an activity which one performs in one’s mind? Why should we think of inner speech as an internalised version of egocentric speech, and not as a different kind of phenomenon which replaces egocentric speech?

The Imagined Speech View and the Actual Speech View can be seen as offering competing answers to these questions. On the Imagined Speech View, internalising egocentric speech would involve beginning to imagine producing this kind of private speech. On this view, the internalisation of egocentric speech would involve a major transformation. An actual phenomenon would be replaced with an imagined one. On the Actual Speech View, however, internalising egocentric speech would just involve beginning to produce this kind of private speech, but doing so silently.

Insofar as the Actual Speech View is to be preferred, we now have a clearer idea of what the ‘internalisation’ of egocentric speech must involve. Vygotsky’s metaphor, that egocentric speech ‘goes underground’, turns out to be fitting. When we begin to produce egocentric speech silently, it becomes unobservable to others, but it does not cease to exist, nor is it replaced with something else.

## **6.2. Other Such Phenomena?**

Generally, one can imagine performing an action which one can actually perform in the external world. It is very natural to think, though, that one cannot actually perform in the mind an action which one can actually perform in the external world. If inner speech is a kind of actual speech, it does not fit this pattern. One can actually speak in the external world and one can imagine speaking – and one can actually speak in one’s mind. This raises the question: are there many other such kinds of action?

One example was mentioned in Section 4.3.2. There is a strong intuition that we can actually perform arithmetic in our minds. This is presumably also the case with respect to many other areas of mathematics: one can actually do algebra, calculus etc. Geometry is an interesting case.<sup>53</sup> Perhaps we have a stronger sense that we imagine the polygons, lines etc. on which we perform manipulations.

---

<sup>53</sup> I refer only to standard Euclidian geometry.

We can modulate the pitch of our inner speech<sup>54</sup> and also the rate at which we produce it. If this is so, then we might think that it is also possible to sing in the mind. This, though, is somewhat more counterintuitive. Perhaps producing actual sounds is essential to singing.

Of course, there are many types of action which can be performed in the external world but not in the mind because their performance presupposes a certain kind of object. You cannot play the flute or plant a tree in your mind; there are no flutes or trees there. However, there are kinds of action where this is not so; these are the kinds of action one might think it is possible to perform in the mind. If one can actually speak in the mind, can one actually gesture? Can one actually laugh in the mind?

However these questions are to be answered, it should inform our understanding of the mind. What is it about some actions which allows us to perform them in the mind as well as in the external world? And what is it about the mind that makes this the case?

### **6.3. Aptness for Moral Judgement**

It would be strange to think that it could be morally wrong or blameworthy to produce a particular inner speech utterance. If one says something rude, or arrogant, or judgemental in inner speech, no moral offence is committed. It would be easy to make sense of this if inner speech were imagined speech, for imagined actions are simply not suitable objects of moral evaluation. But this chapter has sought to show that inner speech is actual speech – and there are many moral theories on which it would be wrong or blameworthy to actually say certain things, e.g. things which are rude, or arrogant, or judgemental. It seems, then, that these theories owe an explanation as to why inner speech utterances are not apt for moral judgement.

It is inevitably put to me when I discuss this issue with others that it would be immoral to produce a racist remark in inner speech. Without seeking to litigate this issue thoroughly, I suspect that people are repulsed by the thought of producing such a remark in inner speech, rather than committed to the view that it is immoral to do so.

---

<sup>54</sup> Recall that this was discussed a little in Section 2.1.

## Chapter Two: Is Inner Speech Dialogic?<sup>55</sup>

### 1. The Problem

There is a theory that inner speech is dialogic. The theory takes as its point of departure Lev Vygotsky's work on the development of inner speech, though Vygotsky himself did not explicitly endorse the view. As the reader will recall from the Introduction, Vygotsky observed in his classic *Thought and Language* (1934/1986) that children first encounter speech in social contexts – e.g. while being spoken to by their caregivers. He thought that, as young children, they then begin to produce a type of external speech – egocentric speech – in which they seem to make remarks about whatever activities are occupying their attention. He suggested that, a little later, this practice is internalised and becomes inner speech. It has been suggested more recently that inner speech not only develops from children's experiences of conversational exchanges but also that it retains the character of these conversational exchanges – i.e. that inner speech itself is dialogic.<sup>56</sup> (On the notion of dialogicality and / or the alleged dialogicality of inner speech, see e.g. Fernyhough 1996, 2004, 2008, 2009; Wertsch 1980, 1991.)

Plainly, however, the theory that inner speech is dialogic seems to run contrary to one very obvious fact about inner speech. Inner speech involves only one participant. There is no distinct interlocutor, as there is when one is engaged in external dialogue. The only person available to respond to a remark one makes in inner speech is oneself.

Is inner speech dialogic? This is the question I will take up in this chapter. I am going to argue against the theory and then briefly suggest an explanation as to why it is that inner speech may often seem to be dialogic, even though it is not. Before beginning, however, there are a number of important preliminary matters to address.

First, an issue about the scope of the theory. The theory is presumably not that inner speech is always dialogic. Take a case in which someone sitting in her office hears a soft knocking sound and produces a single inner speech utterance – perhaps 'Ah, someone at the door!'. It is very hard to see how one could think that this single utterance is dialogic, in any sense. One does find in the literature some comments which seem to suggest the ambitious claim that inner speech is always dialogic – e.g. that inner speech is 'irreducibly dialogic' (Fernyhough

---

<sup>55</sup> The majority of this chapter has been published in the *Journal of Consciousness Studies* (Gregory 2017).

<sup>56</sup> It has also been claimed that this idea is implied in Vygotsky's own work (Fernyhough 2004, p. 53, citing Vygotsky 1934/1987). As foreshadowed, Fernyhough generally uses the 1987 translation of *Thought and Language / Thinking and Speech*; see Footnote 9. As this chapter is concerned primarily with Fernyhough's work, I will do so as well throughout the remainder of the chapter.)

2004, p. 50, citing Fernyhough 1996 and Wertsch 1991), or that a more appropriate term than ‘inner speech’ would be ‘inner dialogue’ (Wertsch 1980, p. 151). Still, I will assume that the theory is that inner speech is at least sometimes dialogic; this interpretation surely places the view in its best light. Maybe some theorists would urge that inner speech is often dialogic – but any such claim will obviously be defeated if it can be shown that inner speech is not even sometimes so.

Next, a note on the relationship between the last chapter and this one. In the last chapter, I sought to show that inner speech is a type of actual speech, not a type of imagined speech. One might think, however, that inner speech could only be dialogic if producing inner speech could involve imagining an interlocutor (or interlocutors) speaking. This would require an extremely broad understanding of inner speech – much broader than the understanding represented by what I called the Imagined Speech View. If one does have this extremely broad understanding of inner speech, then one could develop the theory that inner speech is dialogic in this way. However, I will shortly show that there is also an alternative way the theory can be developed which comports with the view that inner speech is a kind of actual speech. For a reason that will become clear, I think this latter, alternative way of developing the theory is to be preferred – and not just because it is compatible with the view that inner speech is a kind of actual speech. In any case, the argument I will provide for the view that inner speech is not dialogic will have force against the theory whichever way it is developed.

Finally, a word on sources. The theory that inner speech is dialogic has its origins in the writings of some important historical figures including, as well as Vygotsky, the Soviet literary thinker, Mikhail Bakhtin, and the pragmatist philosophers, Charles Sanders Peirce and George Herbert Mead. I will make no attempt at exegesis of the works of these thinkers.<sup>57</sup> It will be convenient at a couple of points to refer to Vygotsky and Bakhtin (including, shortly, in providing some background to the theory) but their work will not be examined in detail. Rather, I am concerned in this chapter with the theory that inner speech is dialogic as it has been developed in contemporary work.

In the second section of this chapter, I will review the theory that inner speech is dialogic, focussing on the way it has been developed by one particular theorist, Charles Fernyhough. In the third section, I will outline and evaluate the arguments which support the theory. In doing so, I will acknowledge that one important part of the theory should be accepted. In the fourth section, I will argue against the notion that inner speech involves a kind of dialogic exchange, as that notion is explicated by Fernyhough. To make this argument, I will borrow some

---

<sup>57</sup> Those interested in the contributions of these thinkers could start with Fernyhough (1996) and Wiley (2016).

extremely influential material from the philosophy of language. In the fifth section, I will respond to some possible objections. In the sixth, I will offer a brief explanation as to why inner speech can seem dialogic, even though it is not. I will conclude this chapter as I concluded the last, by tracing out some implications of the position which I argue for.

## **2. The Dialogue Theory**

The thinker who has done most to develop the theory that inner speech is dialogic is Charles Fernyhough. Accordingly, it will serve to focus on his development of the theory. Fernyhough is a psychologist and some of the terminology he uses in setting out his theory may not be familiar to philosophers, and indeed may seem a little odd. Also, some aspects of his theory permit of different interpretations; one example of this has already been foreshadowed. In order to be faithful to Fernyhough's theory, I will follow his terminology, though seek to clarify it where possible.<sup>58</sup> I will also indicate where parts of the theory appear ambiguous.

There are several aspects to Fernyhough's presentation of the theory, which I will now work through in turn.

### **2.1. From Interpersonal to Intrapersonal: Some Context**

Fernyhough investigates inner speech within an explicitly Vygotskian framework. According to Fernyhough, a part of this framework is the view that the development of every mental function follows a distinctive pattern. He writes:

In his 'general genetic law of cultural development' (Vygotsky, 1931/1997), Vygotsky claimed that every mental function appears twice in development: firstly on the interpsychological plane (that is, as a function distributed between more than one individual), and secondly on the intrapsychological plane, as an internalised version of that previously external function. (Fernyhough 2004, p. 54).

Put aside the question whether this 'general genetic law' is plausible. The relevance for present purposes is that Fernyhough takes the developmental trajectory which Vygotsky attributes to inner speech – i.e. from experience of conversational exchanges, to egocentric speech, to inner speech – to be just a particular instance of this law (Fernyhough 2004, p. 54; see also Vygotsky 1934/1987).

---

<sup>58</sup> Mostly, not a lot turns on the terminology which is not wholly precise, and it is sufficient to operate with an intuitive understanding of it. The key term, which will be introduced shortly, is 'self-regulating', and interpreting that term is central to the project of the chapter.



This context helps to understand what is being claimed by the theory that inner speech is dialogic. The view is not just that inner speech is descended from dialogic exchanges or that it resembles them. We now see why. If one believes that the development of inner speech begins with a child's experience of dialogic exchanges, and if one believes that its developmental trajectory follows the 'general genetic law', then one must also believe that inner speech is actually 'an internalised version' of those dialogic exchanges. Inner speech does undergo some important changes as it is internalised; there will be more on this shortly. Still, what takes place is 'the reconstruction of external dialogue on the internal plane' (Fernyhough 2008, p. 235).

## 2.2. What is Dialogue?

But what exactly is meant by the term, 'dialogue'? In a (2009) paper, Fernyhough outlines five criteria which he holds must be met in order for a 'mental process', such as the generation of inner speech, to be considered dialogic (p. 43).<sup>59</sup> Let us consider each of them, incorporating references to other statements of the view for elaboration.

- 1) 'Simultaneous accommodation of multiple perspectives' (Fernyhough 2009, p. 44).

Fernyhough writes that a perspective is a 'position-bound representation of an orientation toward a state of affairs' (2009, p. 44). He elaborates:

Although some of these orientations may be purely cognitive-perceptual, ('I see that X is behind Y'), others may describe an affective orientation ('I want X'), a situated motivation to act ('I need to move Y to get X'), and so on. ... [E]ach perspective constitutes a description *for an agent in a particular spatiotemporal location*. (Fernyhough 2009, p. 44, emphasis original; see also Fernyhough 2008, p. 234 citing Barresi & Moore 1996, Hobson 1995, Vygotsky 1934/1987).

Elsewhere, he indicates that perspectives can also be 'theoretical' (Fernyhough 1996, p. 51) or 'epistemic' (Fernyhough 2008, p. 234 citing Barresi & Moore 1996).

In philosophical terminology, affective, motivational, theoretical and epistemic perspectives can be understood as sets of propositions to which one takes various attitudes.<sup>60</sup> The attitudes can be either cognitive or conative. The cognitive attitudes need not be beliefs; one might just temporarily entertain a proposition or propositions (Fernyhough 2008, p. 34). A perceptual perspective, by contrast, is a viewpoint in space from which one can, for example, see some

---

<sup>59</sup> An earlier statement of the view, including a discussion of the work of Mikhail Bakhtin from which the account borrows, can be found in Fernyhough (1996).

<sup>60</sup> Thanks to Daniel Stoljar for this suggestion.

physical object. One might actually occupy such a perspective or imagine doing so. There can also be hybrid perspectives where one e.g. imagines seeing something and holds or entertains certain propositional attitudes towards it. A ‘situated motivation to act’, where one (say) sees an object and wants to do something with it, would be an example of a hybrid perspective.

What is it for multiple perspectives to be *accommodated simultaneously*? There are two ways this might be understood. In an early paper, Fernyhough (1996) draws inspiration from the work of Mikhail Bakhtin and writes:

Any utterance can ... manifest more than one perspective ... at the same time. Bakhtin's ([1975/1981]) example of such a ‘multivoiced’ utterance is provided by a sentence from Dickens’ *Little Dorrit* (1857/1967, p. 621):

‘But Mr Tite Barnacle was a buttoned-up man, and *consequently* a weighty one’ [emphasis added by Fernyhough].

The main part of this sentence is expressed in the voice of ‘common opinion’, whereas the word *consequently* appears to represent an interjection or ‘response’ by the voice of the author. (Fernyhough 1996, p. 49).

It is not at all clear to me why one might wish to say that multiple perspectives are expressed simultaneously in the quotation from *Little Dorrit* and not, say, in rapid succession. Still, we can at least make sense of the suggestion that two perspectives are represented within one utterance. This applies as well to utterances in inner speech. Consider the case of a graduate student who is under pressure from her supervisor to finish a paper, though does not herself see the urgency of it. She might produce the inner speech utterance, ‘I need to finish this soon’. It could be that two sets of propositional attitudes are represented by the utterance: those the student believes (‘I need to finish this’) and those she takes her supervisor to believe (‘soon’).

More recently, however, Fernyhough has written that ‘[d]ialogic thinking is made possible when two or more ... perspectives are held in mind at the same time’ (Fernyhough 2009, p. 44). So perhaps it is enough that one produces utterances in inner speech representing various perspectives in turn, at each point holding the others in mind. This notion of holding a perspective in mind is not explained. Presumably, though, it involves storing some representation of it in working memory. In the case just considered, for example, the subject might retain the first part of the utterance (‘I need to finish this’) in working memory while producing the second part (‘soon’), which is associated with a different perspective.

For the purposes of what follows, it will not matter whether one prefers to develop the notion that multiple perspectives can be accommodated simultaneously in one of these ways or the other.

2) ‘Perspectives are represented semiotically’ (Fernyhough 2009, p. 44).

In inner speech dialogue, perspectives are represented by the production of linguistic utterances. However, it appears that there are two ways this apparently straightforward point can be developed.

First, Fernyhough has written that inner dialogue involves ‘virtual interlocutors’ (Fernyhough 2008, p. 241). The idea seems to be that one can represent utterances in inner speech not as issuing from one’s own (perceptual or other) perspective, but as being spoken by an interlocutor who has a different perspective from oneself – i.e. who occupies a different spatiotemporal location vis-à-vis some object, or who holds a different set of beliefs and desires.

This is not to say that subjects need to imagine particular individuals occupying different perspectives. Rather, a virtual interlocutor could be a kind of generic interlocutor without, for example, a distinct voice. I think it is clear that an episode in which one imagines someone other than oneself speaking is not an instance of inner speech. Recall that not even a proponent of what was called the ‘Imagined Speech View’ in Chapter One would claim that it is. This is the point foreshadowed in the first section of this chapter when I noted that one might think that inner speech could only be dialogic if one has an extremely expansive notion of inner speech.

In any case, a passage from the same (2008) paper suggests an alternative way to develop the idea that multiple perspectives are represented by the production of inner speech utterances. Fernyhough describes the allegedly dialogic nature of the egocentric speech produced by a child attempting a jigsaw puzzle in her mother’s presence:

At several places [in an extract from a transcript] the child appears to be asking questions of herself and then answering them. For example, in making the utterance ‘That shouldn’t go there, should it?’, the child adopts an alternative, adult perspective on the task *and represents it for herself in overt speech* while a response can be generated. (Fernyhough 2008, p. 236 referring to Fernyhough 1994, emphasis added).

Perhaps, then, one represents a perspective other than one’s own in inner speech as follows. One temporarily adopts the other perspective, perhaps by imagining having a particular perceptual viewpoint, or by entertaining a particular set of propositional attitudes. One then produces inner speech. Whatever inner speech utterances one produces will be a representation of that perspective. On this interpretation of the suggestion that inner speech involves representing different perspectives, it need not be presupposed that inner speech can involve imagining an interlocutor speaking.

I think the second way of developing the idea is to be preferred, for the following reason. As already noted, the theory under consideration is not merely that episodes of inner speech can resemble dialogue. Rather, the theory is that some episodes of inner speech actually are dialogic. In dialogue, different perspectives actually are occupied. In interpersonal speech, of course, the different perspectives are occupied by different individuals. This is obviously not the case in inner speech. Still, if a subject who is producing inner speech adopts different perspectives – and does not merely imagine an interlocutor doing so – then we can make sense of the idea that a kind of dialogue actually takes place.

In what follows, I will generally assume that, on Fernyhough's theory, a subject represents a perspective in inner speech in the second way outlined – i.e. by adopting the perspective and then producing inner speech. In one instance, I will note that a consideration which might be appealed to in support of the theory only really provides support for the theory interpreted in the first way I outlined, i.e. as holding that inner speech involves virtual interlocutors. As foreshadowed, though, the argument I will provide against the theory that inner speech is dialogic has force against the theory interpreted either way. I will return to this point in Section 4.5.

3) 'Multiple perspectives are flexibly coordinated' (Fernyhough 2009, p. 44).

The perspectives involved in a mental dialogue are not fixed: 'internal dialogue involves the ongoing modification of perspectives by their dialogic counterparts' (2009, p. 44). This is because of a subject's 'ability to treat each perspective as if it had come from another person, and to respond to it accordingly' (2009, p. 44).

Put another way, the idea here is just that perspectives can develop during the course of a dialogue. Take the following case. A subject believes that the weather is very warm but does not currently have a desire to go swimming at a nearby beach. She produces the inner speech utterance, 'It really is hot today'. A moment later, entertaining the relevant belief, she produces the inner speech utterance, 'And it's going to get a lot hotter'. It may be that the desire to go swimming at the nearby beach is then added to the set of beliefs and desires that the subject holds – i.e. to the perspective represented by the original utterance. Her perspective has been modified in the course of the short dialogue, just as it may have been modified if another individual had said to her, 'And it's going to get a lot hotter'.

- 4) ‘Multiple perspectives preserve the triadic intentional relations of external dialogue’ (Fernyhough 2009, p. 44).

Fernyhough describes the ‘intentional relations’ involved in external dialogue as follows: ‘Each agent has a perspective on (a) the ‘object’, or element of reality being jointly attended to ...; and (b) the other agent’s perspective on the object’ (Fernyhough 2008, p. 235). This seems right. In any external dialogue, a speaker must attend not only to the object of the conversation but also to what they take their interlocutor to believe, desire etc. about that object.

Regarding internal dialogue, the requirement is that the same ‘intentional relations’ obtain between the perspectives which are represented as obtain between agents in external dialogue. ‘Each perspective in internal dialogue’, Fernyhough writes, ‘represents a particular orientation to a particular element of reality. In addition, the perspectives accommodated in dialogue bear relations to *each other*’ (Fernyhough 2009, p. 44, emphasis original).

Given that perspectives are best thought of as perceptual viewpoints or sets of propositional attitudes, it is hard to make sense of the notion that perspectives themselves bear intentional relations to one another. It would be better to say that a subject, when representing one perspective, remains cognisant of the other perspectives they may represent – in the same way that we remain cognisant (as much as we can) of the perspectives of our interlocutors in interpersonal dialogue, whatever exactly this might involve. I assume this is what Fernyhough means.

- 5) ‘Mental dialogue is open-ended and self-regulating’ (Fernyhough 2009, p. 44).

There are two points here. What is it for internal dialogue to be open-ended? Fernyhough writes:

In the same way that external dialogue is always open to the alternative perspective of the interlocutor, so the dialogic structuring of thought ensures that there is always an ‘open slot’ available for an alternative perspective. (Fernyhough 2009, pp. 44-45).

The idea seems to be that there is always the possibility of representing alternative perspectives in inner speech. Whatever utterance one might produce in inner speech, one can always produce a further utterance, which represents a different perspective, in response to the first utterance.

What is it for internal dialogue to be self-regulating?

Just as the external dialogue between two individuals requires no direction from a third source, so there is no need to postulate a superordinate supervisory system responsible for directing the flow of internal dialogue. (Fernyhough 2009, p. 45).

External dialogue proceeds as interlocutors respond to one another's utterances. Likewise, inner speech dialogue proceeds as utterances representing different perspectives are produced in response to one another. But that is all there is. We need not suppose there is a 'supervisory system' coordinating the process.<sup>61</sup>

### 2.3. Abbreviation

So, inner speech dialogue is an internal reconstruction of external dialogue. It involves the 'simultaneous accommodation of multiple perspectives', where these perspectives can be perceptual, affective, motivational, epistemic etc. A subject represents a perspective linguistically, by adopting the relevant perspective (where this could mean imagining having a particular perceptual viewpoint, or entertaining a particular set of beliefs and desires) and producing inner speech. Perspectives can develop through the course of inner speech dialogue. They bear 'intentional relations' not only to some object but also to one another (or, rather, a subject, when representing one perspective, remains cognisant of the other perspectives they may represent). The possibility of representing an alternative perspective is always available. Inner speech dialogue does not depend on an exterior system for direction or oversight.

Still, inner speech dialogue differs from external speech dialogue. This is because, in the Vygotskian framework Fernyhough adopts, inner speech undergoes some significant changes during the process of internalisation. Most importantly for present purposes, typical inner speech becomes extremely abbreviated compared with external speech (Fernyhough 2004, p. 54, 2008, 2009; Vygotsky 1934/1987).<sup>62</sup> After all, inner speech is something like the limit case among 'cases of external dialogue where well-established shared assumptions between

---

<sup>61</sup> Insofar as Fernyhough's model of inner speech dialogue requires no 'superordinate supervisory system', he suggests that it avoids 'the problems of infinite regress' which arise for theories positing a 'central supervisor' (Fernyhough 2009, p. 45, citing Dennett 1991 and Fernyhough 1996, 2004). Regress problems, however, may not be so easily escaped. Suppose one suggested that the action of producing inner speech is initiated by the operation of some mental subsystem. What initiates the operation of that mental subsystem?

The argument of this chapter will be that inner speech is not dialogic because episodes of inner speech are not self-regulating in the way that interpersonal dialogues are. In making this argument, I will assume that it does not generate any additional kind of regress problem that a complete account of mental agency would need to deal with.

<sup>62</sup> Saying that inner speech undergoes changes such as increasing abbreviation during the process of internalisation obviously does not contradict the claim I made in Section 6.1 of Chapter One, that inner speech remains a kind of actual speech as it is internalised (assuming Vygotsky is right that inner speech does develop through a process of internalising egocentric speech). There is no inconsistency at all in the notion that inner speech is a silent and abbreviated kind of actual speech.

the interlocutors mean that only minimal overt speech is needed' (Fernyhough 2008, p. 235). Vygotsky believed, for example, that subjects are frequently omitted in inner speech utterances and only predicates expressed (Fernyhough 2004, p. 54, 2008, 2009; Vygotsky 1934/1987).<sup>63</sup> And this 'syntactic abbreviation of inner speech', Fernyhough suggests, 'is responsible for our experiencing it not as a sequence of fully formed utterances but rather as a fragmentary, condensed series of verbal images' (Fernyhough 2004, p. 54).

Fernyhough writes that '[o]ne implication of this extensive abbreviation of inner speech is that mental dialogue develops away from the "give-and-take" patterning of external dialogue, to a situation where multiple perspectives are represented at the same time' (2008, pp. 235-236). So, generally, multiple perspectives will be represented within a single, possibly quite contracted utterance, rather than in a succession of fully-formed comments and replies (though, as we saw, it is not clear that perspectives within a single utterance are represented simultaneously rather than in rapid succession). But inner speech does not always take this truncated form. Especially when subjects confront cognitively demanding situations, such as 'a difficult problem or challenging task', they may produce inner speech which exhibits a 'more explicit dialogic form' (Fernyhough 2009, p. 46; see also Fernyhough 2008) – e.g. series of questions and answers, or assertions and responses.

### **3. Assessing the Case**

It will be convenient to assess the case for the theory that inner speech is dialogic in two parts. I believe the notion that we represent different perspectives in inner speech is true, but that some pieces of evidence that might be cited in its favour are more impressive than others. In a first subsection, I will review the evidence and calibrate its force. I also believe, however, that the notion that a subject produces any kind of dialogue as a consequence of representing different perspectives in inner speech is not strongly supported by the considerations which are most naturally marshalled in its favour. In a second subsection, I will itemise those considerations and also point to their limitations.

#### **3.1. Representing Perspectives**

In reviewing the evidence that we represent different perspectives in inner speech, it will be helpful to bear in mind the two kinds of perspectives that are allegedly represented: perceptual

---

<sup>63</sup> Fernyhough provides a range of references to studies which found evidence of abbreviation in the egocentric or private speech which, on the Vygotskian model, is a developmental predecessor of inner speech: Feigenbaum 1992; Goudena 1992; Pellegrini 1981; Wertsch 1979; Winsler, et al. 2003.

perspectives, and those perspectives which consist in sets of propositional attitudes. The notion that we can represent different perceptual perspectives in inner speech has some introspective appeal. Recall the example that was used in Section 3.4.4 of Chapter One: a young equestrian is thinking about a horse she is considering buying. Visualising the horse from one side, she produces the inner speech utterance, 'He's so beautiful'. I said there that this does not amount to imagining producing private speech in a context other than one's present context. It is not like a case in which the young equestrian imagines actually being in a paddock near the animal. However, insofar as the subject is visualising the horse from one side, there is nonetheless a sense in which she has taken up a perceptual perspective on it. We can also show how the subject might switch perspective, by expanding the case just a little. Immediately after producing the first utterance, she may recall, and visualise, a blemish on the animal's other flank, and produce the utterance, 'Though he does have that mark on his side'.

The experience of entertaining a set of propositional attitudes and producing the kind of inner speech utterances one would produce if one actually held those propositional attitudes is familiar. One might do this while having an argument and trying to anticipate their antagonist's lines of attack.

Secondly, there is some neuropsychological evidence which could be interpreted as supporting the notion that we represent perspectives which consist in sets of propositional attitudes in inner speech. A recent fMRI study reported by Alderson-Day et al. (2016) found that some episodes of inner speech involved increased activation in parts of the brain which are associated with theory of mind reasoning. These were episodes in which subjects were cued to 'generate inner speech' within (imagined) scenarios involving multiple participants – e.g. visiting one's old school and having a conversation with a teacher, or calling a relative and having a conversation with them. It would seem, then, that the subjects were reasoning about the beliefs and desires of the putative other individuals in the scenarios. Presumably, these perspectives were then represented by the subjects in the inner speech they produced.

This evidence likely supports the theory that we represent perspectives of the relevant kind in inner speech only if the theory is interpreted as holding that inner speech can involve 'virtual interlocutors'. For, surely, the most natural way for subjects to respond to the cues in the experiment would have been to imagine having a conversation with a former teacher or a relative. As above, I think there is little to be said in favour of this interpretation. Still, if one does prefer to interpret the theory that inner speech is dialogic in the way which allows for 'virtual interlocutors', then the evidence just reviewed certainly is relevant.

Next, there is some potentially relevant evidence in the ways that we address and refer to ourselves in inner speech. Consider the case of a foreign language student, noting a deficiency



in her skills, who produces the inner speech utterance, ‘You need to revise the adjective endings’. Plausibly, the student is representing her teacher’s beliefs about her own (i.e. the student’s) area of weakness in the subject. Or an orienteer, trying to visualise where she is on a course: ‘You must have taken a wrong turn.’ The practice of producing inner speech utterances addressed to ourselves in this way might be taken as evidence that some inner speech utterances represent perspectives other than our own and from which it makes sense to address ourselves.

Of course, the use of the second-person pronoun or one’s own name in an utterance in inner speech does not entail that the utterance represents a different perspective from one’s own. Take the case in which the foreign language student produces the utterance, ‘You need to revise the adjective endings’. It could be, as suggested above, that the student is representing the beliefs of her teacher. But now suppose that the student herself thinks that she needs to improve her grasp of the adjective endings, notwithstanding her teacher’s reassurances that she does not. Then it would seem that the only possibility is that the utterance represents the student’s own beliefs about her language skills and her desire to improve them – i.e. her own perspective. Still, the use of second-person language to refer to ourselves in inner speech does seem to suggest strongly that, at least sometimes, our utterances represent perspectives other than our own.<sup>64</sup>

Probably, the introspective evidence for the notion that we represent multiple perspectives in inner speech is strong enough to establish the point by itself. It would be very surprising if someone denied that they represent different perspectives in inner speech. It would be very surprising also if someone doubted that the use of second-person references to themselves in their inner speech was at least sometimes a sign of their representing different perspectives. Whether one is moved by the neuropsychological evidence, of course, will depend on how one thinks of inner speech, specifically, whether it can involve ‘virtual interlocutors’.

### **3.2. Representing Perspectives in Dialogue**

But even if it is true that producing inner speech can involve adopting and representing multiple perspectives, is it true that the resulting inner speech is dialogic? That is, should we believe the rest of the theory?

---

<sup>64</sup> One could make a parallel point about cases in which we address ourselves in the third-person. Recall again the anecdote about Jerry Fodor from the Introduction to the thesis: when thinking hard, he produces inner speech such as ‘C’mon, Jerry, you can do it’. Perhaps he is representing his own desire to succeed; then again, he could be representing his Granny’s desire that he do so.

The most obvious reason for doing so again comes from introspection. Very often, series of inner speech utterances resemble lines from a dialogue, especially in the case of what Fernyhough calls expanded inner speech (Fernyhough 2009, p. 46 (and references therein); Voloshinov 1973, p. 38 as quoted by Wertsch 1980, p. 152). Take a case in which someone produces a series of questions and answers as they work through some problem. Perhaps they are reading a detective novel and trying to determine which character is the villain. At one point, they produce the inner speech utterance, ‘Who stood to benefit from the crime?’. Next, as if in response: ‘Well, the chef, the chauffeur and the butler all did.’ Or take a case in which someone is deliberating over options for dinner. At one point, they produce the utterance, ‘That new Chinese place is great’. Then, immediately afterwards: ‘But it’s expensive. Better eat at home tonight.’ Certainly, if the parts were spoken aloud by different individuals in conversation, we would not hesitate to affirm that the exchange satisfied all of Fernyhough’s criteria for dialogue.<sup>65</sup> Fernyhough adds that representations of inner speech in literature ‘often have a clear dialogic structure’ (Fernyhough 2009, p. 46, citing Lodge 1990).

Some empirical evidence has been analysed in the same way. Fernyhough 2009 (p. 46) writes that ‘several studies [report] that children’s private [or egocentric] utterances can take the form of self-answered questions’ (Fernyhough cites Feigenbaum 1992 and Kohlberg et al. 1968 as examples. Supposing, as Vygotsky suggests, that this type of speech is a developmental forerunner to inner speech, one could conclude that such sequences of utterances – with their ‘explicit dialogic structure’ (Fernyhough 2009, p. 46) – are also likely often instantiated in mature inner speech. And, again, we would certainly affirm of such sequences that they were dialogic if the parts were spoken aloud by distinct individuals.

Some further empirical evidence shows that subjects generally conceive of their own inner speech in a way which is consistent with the hypothesis that it is dialogic. In a survey conducted by McCarthy-Jones & Fernyhough (2011), adult respondents were asked to indicate whether certain propositions described their inner speech. In each case, a subject could indicate that a proposition:

- ‘Certainly applies to me’;
- ‘Possibly applies to me’;
- ‘If anything, applies to me slightly’;
- ‘If anything, slightly does not apply to me’;
- ‘Possibly does not apply to me’; or
- ‘Certainly does not apply to me’ (p. 1588).

---

<sup>65</sup> Or, more precisely, that they satisfied versions of Fernyhough’s criteria adapted for interpersonal dialogue (as presented, Fernyhough’s criteria are criteria for internal dialogue).

The propositions which the authors took to be relevant to the supposed dialogicality of inner speech were as follows:

‘When I am talking to myself about things in my mind, it is like I am going back and forward asking myself questions and then answering them’;

‘My thinking in words is more like a dialog with myself, rather than my thoughts in a monolog’;

‘When I am talking to myself about things in my mind, it is like I am having a conversation with myself’; and

‘I talk back and forward to myself in my mind about things’ (McCarthy-Jones & Fernyhough 2011, p. 1589).

In 27.6% of responses to these and similar propositions, subjects indicated that the relevant proposition ‘[c]ertainly applies to me’; in 31.3% of responses, that the relevant proposition ‘[p]ossibly applies to me’; and, in 18.3% of responses, that the relevant proposition, ‘[i]f anything, slightly applies to me’. In a study reported by Alderson-Day et al. (2014), 80.1% of respondents endorsed to some extent a majority of the same items.

How much does all this establish? There are a couple of points to note. First, the notion that we often produce series of utterances in inner speech which resemble external dialogue seems to be well supported. The introspective as well as the empirical evidence all testifies to this. That said, the study by McCarthy-Jones & Fernyhough should probably be approached with a measure of caution. It is true that, in a large majority of responses to the propositions which the authors took to be suggestive of dialogic inner speech, subjects indicated some level of assent (a total of 77.2%). But it was only in 27.6% of responses that subjects indicated that a proposition ‘[c]ertainly applies to me’.<sup>66</sup> It is also true, then, that in a large majority of cases, subjects could not confidently assent to propositions indicating that their inner speech is dialogic.<sup>67</sup> So the support provided by that study, at least, is relatively weak.

Secondly, it is striking that all of these sources are only really relevant to series of utterances in what Fernyhough calls ‘expanded’ inner speech. Consider again the student under pressure from her supervisor to finish a paper who produces the utterance in inner speech: ‘I need to finish this soon’. Fernyhough would take this to be an instance of condensed inner speech insofar as two perspectives are represented in the same utterance, the student’s (‘I need to finish this’) and the supervisor’s (‘soon’). There is no support for the notion that such an utterance has any of the characteristics of dialogue in the finding that children produce egocentric speech with an increasing incidence of, for example, question-answer and

---

<sup>66</sup> Cf. Cho & Wu (2014) discussing another aspect of the same study. There will be more on McCarthy-Jones & Fernyhough’s study in the Appendix.

<sup>67</sup> A corresponding analysis of the study reported in Alderson-Day et al. (2014) is not possible, as data indicating the levels of assent to the relevant propositions is not provided.

assertion-response sequences; or in the finding that subjects assent with various levels of confidence to the propositions listed above from McCarthy-Jones & Fernyhough's (2011) study; or, I take, in the introspective evidence. (The same, of course, is true of the external speech utterance, '[b]ut Mr Tite Barnacle was a buttoned-up man, and consequently a weighty one'.)

The evidence (reviewed in the last subsection) that we represent multiple perspectives in inner speech is convincing. The evidence that the resulting inner speech is dialogic has some weight, but hardly establishes the claim. What I wish to suggest now is that this second part of the theory that inner speech is dialogic is vulnerable to a very serious problem.

#### 4. Against the Dialogue Theory

One of the characteristics of inner speech dialogue that Fernyhough stipulates is that it is self-regulating: it proceeds without the coordination of any 'supervisory system'. This is an essential part of the theory that inner speech is dialogic. For the theory holds that inner speech dialogue is an internal version of interpersonal dialogue: inner speech dialogue is not just a developmental descendant of interpersonal dialogue; it is an internal reconstruction of it. And, clearly, interpersonal dialogue is self-regulating; it proceeds without the direction of a 'supervisory system'. So, unless the episodes of inner speech which an adherent to Fernyhough's theory would label dialogic are self-regulating, they cannot actually be dialogic. It is this crucial notion that some episodes of inner speech are self-regulating, however, which I wish to deny.

The argument of this section, then, can be set out as follows:

1. If inner speech is dialogic, then some episodes of inner speech are self-regulating.<sup>68</sup>
2. It is not the case that some episodes of inner speech are self-regulating.
3. Inner speech is not dialogic.

What exactly is meant by 'self-regulating'? What does it mean to say that inner speech dialogue, like external speech dialogue, proceeds without a 'supervisory system' coordinating the process? If inner speech dialogue is an internal version of interpersonal dialogue, then we can suppose that inner speech dialogue must be self-regulating *in the same way* that interpersonal dialogue is self-regulating. And there is, of course, some extremely important

---

<sup>68</sup> '[S]ome episodes of inner speech' should be taken to refer to those episodes which an adherent to Fernyhough's theory would consider dialogic (i.e. it must be the right episodes which are self-regulating).

theoretical work investigating what makes external dialogue possible – how it is that dialogue between multiple individuals can proceed without there being some supervisory system.

In the next subsection, I will review H. P. Grice's famous treatment of conversation and show how it provides an account as to how 'talk exchanges' are self-regulating. I will then demonstrate that this account is not transferable to inner speech. Next, I will show that the same is true of some modified versions of Grice's account and, indeed, of any plausible account of dialogue one might prefer. Finally, as foreshadowed, I will show that the argumentation of this section applies *even if* one thinks that producing inner speech can involve imagining 'virtual interlocutors' speaking (i.e. as opposed to supposing that one represents perspectives in inner speech by adopting those perspectives and producing inner speech from them).

Before beginning, a note on terminology. In much of the literature considered below, the term, 'conversation', is used, rather than 'dialogue', e.g. in Grice's famous account. Plausibly, the terms have slightly different connotations. What matters for present purposes, though, is how a series of utterances exchanged between individuals, or which represent different perspectives, can be self-regulating. Whether one finds an account of this in a discussion about dialogue, or in a discussion about conversation, is not of importance, even if the terms do have some different connotations.

#### **4.1. Grice and Cooperation**

In 'Logic and Conversation' (1975/1989), Grice explores the way that language is used in conversations or 'talk exchanges' (*passim*). Central to his account is what he calls the 'Cooperative Principle' (*passim*). Grice notes that conversations 'do not normally consist of a succession of disconnected remarks, and would not be rational if they did' (1975/1989, p. 26). Rather, they are 'characteristically ... cooperative efforts' in which participants pursue a 'common purpose or set of purposes' or recognise a 'mutually accepted direction' (1975/1989, p. 26). Contributions to a conversation are appropriate only insofar as they serve the common purpose or facilitate progress in the mutually accepted direction. The Cooperative Principle, then, is the following:

Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged. (1975/1989, p. 26).

Assuming such a principle applies, Grice suggests we can identify a number of more specific maxims and submaxims which guide participants in conversations. The maxims and submaxims, which Grice organises into four categories, are as follows:

‘Quantity’:

Maxim 1: ‘Make your contribution as informative as is required (for the current purposes of the exchange).’

Maxim 2: ‘Do not make your contribution more informative than is required.’

‘Quality’:

Maxim: ‘Try to make your contribution one that is true.’

Submaxim 1: ‘Do not say what you believe to be false.’

Submaxim 2: ‘Do not say that for which you lack adequate evidence.’

‘Relation’:

Maxim: ‘Be relevant.’

‘Manner’:

Maxim: ‘Be perspicuous.’

Submaxim 1: ‘Avoid obscurity of expression.’

Submaxim 2: ‘Avoid ambiguity.’

Submaxim 3: ‘Be brief (avoid unnecessary prolixity).’

Submaxim 4: ‘Be orderly.’ (Quotations from Grice 1975/1989, pp. 26-27).

The list is not intended to be exhaustive. Broadly, though, the maxims are rules ‘specially connected ... with the particular purposes that talk (and, so, talk exchange) is adapted to serve and is primarily employed to serve’ (1975/1989, p. 28). In ‘Logic and Conversation’, Grice ‘state[s] [his] maxims as if this purpose were a maximally effective exchange of information’ (1975/1989, p. 28). He does note, though, that talk exchanges serve purposes other than information exchange; this is presumably one reason it might be necessary to add to the list of maxims.

Grice emphasises that the conversational maxims have analogues in other types of cooperative human transactions. His own examples assist in demonstrating this point. If one were fixing a car and asked a companion to pass four screws, it would only be appropriate for the companion to pass four screws rather than two or six (cf. the maxims of Quantity). If one were preparing to make a cake and needed sugar, it would not be appropriate for a companion to provide salt (cf. the maxim of Quality) or to hand the cook a good book while she was trying to mix the ingredients (cf. the maxim of Relation). And, in any cooperative activity, one will ‘expect a partner to make it clear what contribution he is making, and to execute his performance with reasonable dispatch’ (cf. the maxims of Manner) (1975/1989, p. 28).

Grice takes the analogies between the conversational maxims and the expectations which apply in other cooperative human transactions to be

relevant to ... a fundamental question about the Cooperative Principle and its attendant maxims, namely, what the basis is for the assumption which we seem to make ... that talkers will in general (*ceteris paribus* and in the absence of indications to the contrary) proceed in the manner that these principles prescribe. (1975/1989, p. 28).

As Grice notes, we could satisfy ourselves with the empirical fact that people simply do seem to observe the conversational maxims. Grice expresses optimism, however, that some rational basis for observing the maxims might be found. He writes that he

would like to be able to show that observance of the Cooperative Principle and maxims is reasonable (rational) along the following lines: that anyone who cares about the goals that are central to conversation/communication ... must be expected to have an interest, given suitable circumstances, in participating in talk exchanges that will be profitable only on the assumption that they are conducted in general accordance with the Cooperative Principle and the maxims. (1975/1989, pp. 29-30).

He acknowledges that he cannot provide a compelling demonstration, but what Grice is suggesting is that anyone committed to the goals of conversation must design their contributions – and expect their interlocutors to design their contributions – according to principles which apply to cooperative activities generally; hence the importance of the analogies. That is, Grice is suggesting that conversation is an essentially cooperative activity.

Of course, not every utterance needs to conform to the maxims. There can also be ‘conversational implicatures’, in which a speaker communicates something precisely by ‘*flout[ing]* a maxim’, i.e. ‘blatantly fail[ing] to fulfil it’ (1975/1989, p. 30; emphasis original). If a speaker says something which clearly does not conform to one or more of the maxims, but circumstances are such that the speaker can still be presumed to be adhering to the Cooperative Principle, then the speaker should be taken to be implicating something other than what they are literally saying. This is what happens in many instances of irony, sarcasm, etc. By contrast, if a speaker ‘quietly and unostentatiously *violate[s]* a maxim’, then they exploit their interlocutor’s (incorrect) assumption that they are being cooperative and ‘will be liable to mislead’ the interlocutor (1975/1989, p. 30; emphasis original).

It is important to understand the role the Cooperative Principle and the maxims are supposed to play. A listener can make sense of a speaker’s remarks only insofar as the listener can assume of the speaker that they are adhering to the Cooperative Principle. For how can a listener interpret a remark if the remark cannot be supposed to be appropriate to the accepted purpose or direction of the exchange? Indeed, how can a listener know that a speaker means anything at all by producing some series of sounds if that series of sounds cannot be supposed to be appropriate to the accepted purpose or direction of the exchange? Moreover, the presumption that their remarks will be interpreted in line with the Cooperative Principle also guides a speaker in designing their remarks. Observing the Cooperative Principle entails

adhering to whatever subservient maxims are appropriate to the purpose of the exchange, unless one is flouting them in order to generate conversational implicatures.

This is how the role of the Cooperative Principle and the maxims is standardly understood, but there is something we can add. Grice's comment that conversations 'do not normally consist of a succession of disconnected remarks' suggests that they do sometimes consist of just this. Two or more individuals can produce a series of remarks without any regard to what the other(s) are saying, and there is no reason we should not call this a 'conversation'. It is presumed adherence to the Cooperative Principle and the maxims, however, which allows participants in a conversation to make progress with a 'common purpose' or in a 'shared direction'. Moreover, though there is no reason for Grice to make it explicit, it is presumed adherence to the Cooperative Principle and the maxims which allows participants in a conversation or a dialogue to make such progress without the direction of a third-party supervisor. It is this apparatus, we might say, that makes a talk exchange 'self-regulating'.

#### **4.2. Cooperation and Inner Speech**

What about inner speech? Can an episode of inner speech be self-regulating in this way? The answer is no. A series of utterances in inner speech cannot be a cooperative effort because there are no distinct agents who can be said to be cooperating. Accordingly, there is no sense in which one can assume one's interlocutors are adhering to the Cooperative Principle and any maxims derived from it. This is so even supposing that multiple perspectives are represented in inner speech – for it is still the case that there is only one agent producing the utterances which represent the different perspectives.

It is probably obvious that an episode of inner speech cannot be a cooperative effort, but it will still be worthwhile to note a way it can be demonstrated. Suppose someone produces an inner speech utterance which violates the first submaxim of Quality: they produce an inner speech utterance which they know to be false. Producing this inner speech utterance will not bring about the distinctive effect which can be achieved if one does not observe the submaxim, '[t]ry to make your contribution one that is true', when speaking aloud to someone else. There is no one who is liable to be deceived in inner speech, no naïve interlocutor who might form a false belief. But if this is so, then nor can we make sense of an episode of inner speech as a cooperative effort. For there is also no one who is in a position to (say) exchange information with the subject – even if the maxims are observed meticulously.

Now, it might be said that the first submaxim of Quality would never apply to a 'dialogue' in inner speech. Grice's maxims apply on the assumption that the purpose of a particular talk



exchange is ‘a maximally effective exchange of information’, and this would never be the purpose of producing utterances in inner speech. There would be, after all, only one agent involved. But this is just the point. It is true that the purpose of an episode of inner speech would never be ‘a maximally effective exchange of information’ – or anything else which requires coordination between individuals. It is for this reason that an episode of inner speech is not the kind of event to which rules facilitating cooperation can apply. And this, of course, is just to say that an episode of inner speech is not self-regulating in the way that an interpersonal dialogue is.

It might also be said that Grice’s apparatus does not apply to purported dialogues in inner speech only because it is unnecessary. One does not need (for example) maxims exhorting truthfulness in inner speech because one knows whether one is being truthful or not. Again, though, this is really just another way of getting to the central point. Producing series of inner speech utterances is simply not a cooperative effort. If it were, then a set of rules regulating the cooperation *would* be necessary.

### **4.3. Developments in Gricean Theory**

Grice’s theory has been hugely influential: in broad outline, it is accepted as near to universally as anything in analytic philosophy. There have, however, been some interesting suggestions as to how the account might be amended to deal with particular problems or otherwise improved. It will be worthwhile reviewing some of these suggestions, in case some modified version of Grice’s account might be transferable to inner speech, even if the original is not.

#### **4.3.1. Martinich and Authenticity**

Discussing the submaxims of Quality, A. P. Martinich complains that ‘[t]hey suffer from the typical philosophical disease of fixating on serious factual statements, when people often do other things with language’ (1980, p. 219). In their place, he recommends the following:

Be authentic. That is, do not knowingly participate in a speech act for which the conditions for its successful and non-defective performance are not satisfied. (1980, p. 220).

This enjoins speakers not just to assert only what they have reason to believe, but also ‘to promise to do only what [they] can do; apologize only for offences actually given; forgive only persons who have injured [them]’ (1980, p. 220).

Martinich's objection is perhaps not entirely fair: as noted, Grice's schema of maxims is not intended to be exhaustive, and he may well have approved of exactly the kind of submaxim Martinich proposes.<sup>69</sup> In any case, even if Martinich's amendment does better equip the Gricean apparatus to deal with the full range of interpersonal talk exchanges, one can see that it will still not be transferable to episodes of inner speech. For the modified account still supposes that conversation is essentially a cooperative effort. And, even if one makes it explicit that partners in interpersonal dialogue can cooperate on a wider range of projects than just the effective exchange of information, a series of inner speech utterances still cannot be a cooperative effort. This analysis, of course, will apply to any account which supposes that conversation is cooperative.<sup>70</sup>

#### 4.3.2. Kasher and Rationality

But what if one rejects the Cooperative Principle? Asa Kasher (1976, 1982) does just this. Specifically, he doubts that every conversation actually has an 'accepted purpose or direction'. He asks the reader to 'imagine a prolonged and variegated conversation between Socrates and Crito, in which each attempts to convince the other that he is in error' (1976, p. 202). In such circumstances, Kasher suggests, 'the conversants will not have mutual global aims and generally no mutual local aims' (1976, p. 202). It is true that they may share '[t]rivial general aims', e.g. 'to exchange opinions of some sort', but Kasher claims that these 'are not sufficiently specific to have them serve as assayers of the contributions to the conversation, as is required by the cooperation principle' (1976, p. 202).

So how are conversations organised, according to Kasher? In place of the Cooperative Principle, Kasher proposes a much broader 'rationality principle' (1976, p. 205; 1982):

Given a desired end, one is to choose that action which most effectively, and at least cost, attains that end, *ceteris paribus*. (1976, p. 205; also in 1982, p. 32).

It seems plausible that one can generally assume of others that they will adhere to this principle in performing any action.

Now, in some circumstances, the most effective and efficient way to pursue a particular end is to engage in conversation with another person, e.g. if one wishes to inform or influence that other person. And, Kasher believes, the most effective and efficient way to engage in conversation will actually be to adhere to Grice's original maxims. One should, for example,

---

<sup>69</sup> Presumably in addition to the submaxims of Quality, rather than instead of them.

<sup>70</sup> Martinich makes some other suggestions as to how the maxims might be amended, but these other suggestions would not bring about any fundamental changes to Grice's schema.

say as much as, but no more than, is necessary to convey a particular message (cf. the maxim of Quantity); for otherwise one will either fail to convey the message or use excessive effort in doing so. And one should speak sufficiently clearly as to ensure that one's message is successfully conveyed (cf. the maxim of Manner).<sup>71</sup>

So, for Kasher, even if one rejects the Cooperative Principle, the rest of Grice's apparatus is unavoidable. From here, the story is familiar. When a speaker appears to deviate from a maxim but can still be presumed to be adhering to the rationality principle, an implicature will be generated. Moreover, one can see how the operation of the apparatus would make interpersonal dialogue self-regulating. It will be possible for interlocutors to engage in dialogue without the direction of any supervisory system just insofar as they are able to generally assume of each other that they are adhering to the rationality principle and the maxims. But the apparatus is still not transferable to inner speech. We can see this (again) by considering what happens if a subject produces an utterance in inner speech which violates one of the maxims. There is still no one who is liable to suffer any kind of disadvantage in inner speech; nor is there anyone of whom it can be supposed that they are adhering to the maxims.

#### **4.3.3. Sperber & Wilson and Relevance**

One final, broadly Gricean account that should be considered is Dan Sperber & Deirdre Wilson's (1995) relevance theory. Sperber & Wilson explicitly base their theory of communication on a particular conception of human cognition. Could it be that inner speech 'dialogues' are self-regulating in some way that relevance theory explains?

Sperber & Wilson's presentation of their theory is detailed and complicated; here, a cursory summary will have to suffice. For Sperber & Wilson, '[h]uman cognitive processes ... are geared to achieving the greatest possible cognitive effect for the smallest possible processing effort' (1995, p. vii). Human cognitive processes are geared, that is, to efficiently extracting useful information from environments. When we communicate, we take advantage of this. For communicating with someone involves somehow altering their environment – presenting some stimulus from which they can draw inferences.

What guides us in deciding which stimuli we present to others and in interpreting the stimuli others present to us? Sperber & Wilson believe this can be explained with reference to a single principle, rather than a long series of maxims. To state their proposal, it is necessary to

---

<sup>71</sup> The attempt to derive the maxims of Quality and Relation from the rationality principle is somewhat less convincing, but see Kasher 1976, pp. 206-207 for the details; see also Kasher 1982.

introduce some technical terminology. An 'act of ostensive communication' is any act in which one makes it apparent to someone else that one is trying to make something apparent to them. This includes e.g. when one shows someone something: in (say) pointing something out, one conveys to their audience that one is trying to show them something. It also includes when we speak to one another: speaking to someone makes it apparent to them that one is trying to communicate something to them. A stimulus is 'relevant' insofar as it allows a subject to draw useful inferences when combined with background information which is available to the subject, and insofar as the processing cost of doing so is low. The 'presumption of optimal relevance' is a twofold presumption: 1) that information which a communicator is trying to convey to an addressee is sufficiently relevant that it is worthwhile for the addressee to process the stimulus by which it is being conveyed; and 2) that the stimulus being used is the most relevant one the communicator could have used (i.e. will place the lowest possible processing cost on the addressee).

It is now possible to introduce Sperber & Wilson's central proposal, what they call the 'principle of relevance':

Every act of ostensive communication communicates the presumption of its own optimal relevance. (1995, p. 158).

In the context of verbal conversation, the principle operates as follows. A speaker can suppose that when they say something to someone else, this other person will interpret the utterance on the assumption that the utterance is optimally relevant. The very act of speaking operates to attract an addressee's attention and, in so doing, does in fact convey that what the speaker is saying is optimally relevant – i.e. worth the addressee's attention and conveyed in the most relevant way possible. The addressee will then seek to interpret the utterance on the basis of the same assumption. That is, the addressee will interpret the utterance assuming that the speaker believes it worth the addressee's effort to process, but has made it as easy as possible for the addressee to do so. In practical terms, this will mean first considering obvious, literal interpretations of the remark, and resorting to less-obvious, figurative interpretations only if the obvious, literal interpretations are unsatisfactory (i.e. are not such that the speaker might consider them worth the audience's attention).

One can see how this framework, if it is correct, explains how interpersonal conversation is self-regulating. Insofar as each utterance communicates the presumption of its own optimal relevance, it generates an expectation on the part of the listener that it is worth attending to and processing. But what happens if an utterance disappoints this expectation? That is, what happens if an utterance proves not to be optimally relevant? The listener will presumably become sceptical about the value of attending to any ensuing utterance made by the same speaker. But such a situation cannot be sustained for long. Beyond a certain point, a listener

will simply stop attending to a speaker's utterances; the dialogue will break down. If this is to be avoided, speakers need to try to ensure that their utterances are optimally relevant. This is why no third-party supervisor is required to direct the course of a dialogue. A dialogue will proceed only as long as it suits the participants, and only as long as the participants satisfy the expectations of relevance that they generate whenever they speak.

Insofar as this theory of communication is grounded in a theory of cognition, one might suspect that it would be better placed to accommodate inner speech 'dialogue' than the theories considered above. But this is not the case. As we just saw, what makes interpersonal dialogue self-regulating (on relevance theory's terms) is that speakers must generally satisfy the expectations of relevance which they generate automatically when they speak. But no such expectations are generated in inner speech. There is simply no one in inner speech to hold such an expectation.<sup>72</sup>

So, once more, we have an account which plausibly explains how interpersonal dialogue is self-regulating, but it cannot be applied to inner speech. There is no room for the account to do its work – nor any need.

#### **4.4. An Analogy**

We could continue reviewing accounts which explain how self-regulating dialogue between individuals is possible and asking whether these accounts might also apply to episodes of inner speech, but a clear enough pattern has emerged. What makes self-regulating dialogue between individuals possible is the application of some framework which guides speakers in producing utterances and listeners in interpreting them. For even if dialogue is not a cooperative effort, it is still a kind of multilateral effort, and participants' contributions must be coordinated somehow. But a subject producing inner speech depends on no such framework.

At this point, an analogy may assist. Suppose one wanted to play a game of chess with oneself. Would this be possible? One could alternately move the white and black pieces in accordance with the rules, but there would be a problem. When making a move for white, one would not cease to know what strategy one was seeking to execute for black, and vice-versa. The game might closely resemble the familiar game of chess, but it would be different in important respects. One would not actually have an opponent; one could scarcely celebrate winning or regret losing. The game would not exactly be a game of chess. Moreover, no matter how much

---

<sup>72</sup> Knappik (forthcoming) includes a somewhat similar discussion (Section 4).

one modified the rules of chess, the gap between the multiplayer game and the single-player simulation of it would not get any smaller.

Similarly, it may well be that one can adopt multiple perspectives and produce inner speech utterances appropriate to those perspectives. But while (say) entertaining a particular set of beliefs, one would not cease to know what one was thinking while entertaining some other set of beliefs moments before. The series of utterances might closely resemble a series of utterances which a group of individuals could produce in dialogue. But the episode of inner speech would nonetheless be different from the genuine dialogue produced by the group. Whatever is the framework that makes self-regulating dialogue possible, it would not apply to this episode. For even if one does represent multiple perspectives in inner speech, there are no distinct agents whose contributions must somehow be coordinated. Moreover, no matter how much one modifies or redescribes the framework which makes self-regulating dialogue possible, the gap between interpersonal dialogue and episodes of inner speech involving multiple perspectives will not get any smaller.<sup>73</sup>

#### **4.5. Virtual Interlocutors**

I have noted a couple of times that there are two ways that the suggestion that one can represent multiple perspectives in inner speech can be interpreted, depending on one's background commitments about the nature of inner speech. The interpretation which, I suggest, should be preferred is that we represent perspectives other than our own by adopting those perspectives – e.g. by imagining having a particular perceptual viewpoint, or by entertaining a set of propositional attitudes – and producing inner speech. If one believes that producing inner speech can involve imagining someone else speaking, however, then one might think that we can represent perspectives other than our own by imagining interlocutors speaking from those perspectives. Although I have been proceeding on the basis that the former interpretation is best, I foreshadowed that I would show how the argument against the theory that inner speech is dialogic applies also if one thinks that inner speech can involve what Fernyhough calls 'virtual interlocutors'. The reason for this is now relatively easy to see.

The lesson from Grice is that conversation requires some framework according to which participants can 1) interpret their interlocutors' utterances, and 2) design their own contributions so that their interlocutors can interpret them. According to Kasher, this is so even if one rejects Grice's view that conversation is a kind of cooperative endeavour and

---

<sup>73</sup> Of course, Wittgenstein also famously draws an analogy between language and chess (e.g. 1953/2001). I will not take up that connection here.

thinks of it only as a kind of multilateral endeavour. But an exchange with an imagined interlocutor in inner speech – if inner speech can involve an imagined interlocutor – does not involve any such framework. There is no sense in which one cooperates with, or even coordinates one's actions with, an imagined interlocutor. To state the obvious: an imagined interlocutor is not a distinct agent. This is much the same point as that made when we were assuming that representing a perspective just involves adopting it and producing inner speech: doing so does not require a framework for cooperation or coordination, as interpersonal dialogue does. Finally, this is so no matter how exactly one thinks it is that we do cooperate, or coordinate our actions, with real interlocutors.

## 5. Some Objections

There are a few objections that might be raised against the argumentation of Section 4. It will be worthwhile to consider these objections, and to see how they can be met.

### 5.1. Inner Speech and Self-Deception

The thought that there is no one who is liable to be misled in inner speech played an important role in the argumentation of Section 4. It was appealed to in order to demonstrate the point that the principles which one might think of as regulating interpersonal dialogue have no role to play in inner speech. The first objection challenges this assumption. It asks: is it true that no one is liable to be misled in inner speech? Is there not such a thing as self-deception?<sup>74</sup>

There is such a thing as self-deception but it may not be the kind of phenomenon that its name most readily suggests. On one very plausible view on the nature of self-deception developed by Alfred R. Mele (2001), self-deception is not analogous to interpersonal deception. It does not involve some party *A* who believes *p* trying to bring it about that a party *B* believes  $\sim p$ , where *A* is actually identical to *B*. Rather, self-deception occurs when desires (and possibly emotions) lead to the biased collecting or interpreting of evidence. For example, a college student who is romantically interested in a classmate may interpret her refusal to date him as evidence that she is also interested in him and merely feigning indifference. Or a volunteer for a political campaign who greatly admires the candidate she is assisting, but who has been hearing rumours of gender-based prejudice on his part, might seek out evidence confirming her existing opinion of him (Mele's examples, 2001, pp. 26-27).

---

<sup>74</sup> Thanks to Charles Fernyhough and an anonymous referee at the *Journal of Consciousness Studies* for this point.

This is not the place for an extensive discussion of self-deception. For present purposes, it will suffice to note that one can consistently believe in the possibility of self-deception and deny that inner speech is dialogic. Moreover, this is no ad hoc manoeuvre. The characterisation of self-deception which it depends upon has been argued for independently in the relevant literature.

## 5.2. Self-Knowledge

The second objection challenges an assumption which is implicit in the argumentation of Section 4. The complaint is that the idea that there is no one in inner speech who is liable to be misled assumes that we have more direct access to our mental states than may actually be the case. It assumes, for example, that a subject who believes some proposition,  $p$ , and who produces an utterance in inner speech expressing  $\sim p$ , can know what she believes independently of the inner speech that she produces. For some (e.g. Carruthers 2009, 2011; Cassam 2014) have suggested that we only come to know what we believe by interpreting evidence such as the inner speech we produce.<sup>75</sup> And if a subject does not know that she believes that  $p$  when she produces an inner speech utterance expressing  $\sim p$ , then she could plausibly mislead herself about her own beliefs.

There are two ways this objection can be dealt with. First, one could simply observe that, if the argumentation of Section 4 assumes that we can have access to our mental states independently of the inner speech we produce, the assumption is a natural descendant of an assumption implicit in the relevant accounts of interpersonal dialogue and conversation: that we can have access to our mental states independently of the external speech we produce. For how could one ever know that an utterance that they had produced (or were about to produce) aloud was one which they took to be true, authentic or relevant, unless they did have independent access to their beliefs?

But things are a little more complicated than this. In Chapter Three, I am going to develop a theory which does give a role to inner speech in acquiring self-knowledge (though without actually endorsing that theory). In Section 7 of that chapter, I will show that the theory is consistent with the point that it is not possible to deceive oneself by producing a false utterance in inner speech. This is the second way to deal with the objection: showing that one can both attribute a role to inner speech in acquiring self-knowledge and accept the argumentation of this chapter.

---

<sup>75</sup> Thanks to an anonymous referee at the *Journal of Consciousness Studies* for this point.



So the situation is this. Insofar as the assumption made in this chapter is shared by the leading theories of interpersonal dialogue, one cannot reasonably complain that it is relied upon in making an argument against the possibility of inner speech dialogue. Even if one rejects that assumption, however, one can still accept the argument of this chapter – though it will be some time before we see exactly why this is so.

### 5.3. Dialogue without Rules

A third possible objection questions the supposition that the apparatus that makes self-regulating dialogue between individuals possible – whatever exactly that apparatus may be – is necessary for dialogue in inner speech. As was noted early in this chapter, the theory that inner speech is dialogic emphasises the point that inner speech can be thought of as the limit case among ‘cases of external dialogue where well-established shared assumptions between the interlocutors mean that only minimal overt speech is needed’ (Ferryhough 2008, p. 235). Now, suppose two people hold *exactly* the same set of beliefs and desires. Plausibly, they could carry on a dialogue without relying on Grice’s Cooperative Principle and maxims, or indeed on any other such framework. Neither individual would have to worry about the other being able to interpret their utterances: by hypothesis, the two of them would have identical beliefs about the meaning of each utterance. And if this is plausible, then could a single individual not carry on a dialogue in inner speech without relying on Grice’s or any other framework?<sup>76</sup>

There is something odd about the idea of two people with identical beliefs and desires carrying on a dialogue – after all, it is not as if either could tell the other anything new – but the scenario is certainly coherent. However, it should not give a proponent of the view that inner speech is dialogic any comfort. It is essential to the view that inner speech is dialogic that the dialogue involves *different* perspectives, and two people who hold exactly the same beliefs and desires would actually have the same perspective. Even if it would be possible for two people who have the same perspective to carry out a dialogue without relying on e.g. Grice’s framework, this does not show that there can be a dialogue when one person represents multiple perspectives in inner speech.

---

<sup>76</sup> This was suggested to me by David Chalmers, if I have understood him correctly.

## **6. An Alternative View**

What still needs to be done is to offer some explanation as to why inner speech may so often seem to be dialogic, if this is not actually the case. For it certainly is a familiar experience to produce in inner speech series of questions and answers, or assertions and responses etc. As indicated in Section 3.2, I think that the relevant evidence only supports the notion that it is series of utterances such as these – episodes of what Fernyhough calls ‘expanded’ inner speech – which resemble interpersonal dialogue. Recall: an utterance such as the student’s ‘I need to finish this soon’ does not look in any way dialogic, even supposing that multiple perspectives are represented in it. Still, we are left with the question: why should some series of inner speech utterances resemble dialogue, if they are not in fact dialogic?

I suggest that it is a simple illusion. There is good reason to believe that we do represent different perspectives in inner speech. The evidence for this was reviewed in Section 3.1. Now, suppose that, in some situation, you represent a particular perspective in inner speech. You imagine having a particular perceptual viewpoint, or entertain some set of propositional attitudes, and produce an assertion from that perspective about some subject matter. Then, immediately afterwards, you adopt a different perspective – you imagine having a different perceptual viewpoint, or entertain a different set of propositional attitudes – and produce another assertion about the same subject matter. What will this look like? As a series of assertions about the same subject matter produced from different perspectives, it will resemble an interpersonal dialogue. But this is just because you have adopted different perspectives. It has nothing to do with the framework – whatever exactly that framework is – that coordinates interpersonal dialogue. The same will be true of any series of utterances which one produces successively from two or more perspectives. The resemblance to interpersonal dialogue will be strongest when one produces a series of utterances alternately representing a small number of different perspectives in turns. Still, such a series of utterances will not be dialogic; it will just look like it.

## **7. Some Implications**

My project in this chapter has been to demonstrate that inner speech is not dialogic, in the way that Charles Fernyhough explicates the notion, because it is not self-regulating. To do this, I have reviewed H. P. Grice’s leading account of conversation; shown how this account can explain why interpersonal dialogue is self-regulating; and shown that the account is not transferable to inner speech. I have also shown that this will be true of any account which plausibly explains why interpersonal dialogue is self-regulating.

I close this chapter in the same way as Chapter One, by exploring some implications of the position I have argued for.

### 7.1. Conditionals

The relationship between the natural language, indicative conditional and the material conditional has been a topic of major controversy in analytic philosophy. In setting out the problem here, I assume familiarity with elementary logic. The natural language, logical connectives – conjunction, negation, and (both inclusive and exclusive) disjunction – are truth-functional. The truth or falsity of sentences containing only sub-sentences and these connectives is determined entirely by the truth or falsity of the sub-sentences. It is tempting to think that the natural language indicative conditional is also truth functional. That is, it is tempting to think that the truth or falsity of a sentence containing only sub-sentences and the natural language indicative conditional (with or without any of the other natural language connectives) is also determined entirely by the truth or falsity of the sub-sentences. Motivation for this view comes from the fact that  $A \rightarrow B$  (i.e. if A then B, where  $\rightarrow$  signifies the indicative conditional) seems to be equivalent to both  $\sim A \vee B$  (i.e. not A or B) and  $\sim(A \& \sim B)$  (i.e. it is not the case that both A is true and B is not true), both of which do contain only sub-sentences and truth-functional connectives.

However, the best truth-functional model of the natural language indicative conditional which we can provide is that of the material conditional, and treating the indicative conditional as equivalent to the material conditional has some odd consequences. For example, take a natural language sentence containing a conditional as the main connective, a false antecedent and a true consequent, e.g. if Marseille is the capital of France then Berlin is the capital of Germany. Supposing the indicative conditional is equivalent to the material conditional, such a sentence would be true, though this does not seem at all like a natural thing to say. Likewise, take a natural language sentence containing a conditional as the main connective, a false antecedent and a false consequent, e.g. if King George VI was a President of the United States then Franklin Roosevelt was a King of England. Again, such a sentence would be true, though this also does not seem like a natural thing to say.

Grice believed that the indicative conditional is equivalent to the material conditional – i.e. that the indicative conditional is truth-functional. Use of the conditional in a conversational context, though, like any other piece of language, will be subject to the Cooperative Principle and the maxims. One point Grice makes is that asserting certain conditionals will ‘offend against the first maxim of Quantity’, notwithstanding that they may be true (Grice 1989, p. 61). We can see the point making use of the examples from the last paragraph. Suppose a

speaker knows that Marseille is not the capital of France but does not know whether Berlin is the capital of Germany. Then it will be both more informative and briefer simply to say ‘Marseille is not the capital of France’ than to say ‘if Marseille is the capital of France then Berlin is the capital of Germany’ (notwithstanding that the speaker would take this conditional to be true, insofar as they know that the antecedent is false). Or suppose a speaker knows that King George VI was not a President of the United States and that Franklin Roosevelt was not a King of England. Then it would be more informative to say ‘King George VI was not a President of the United States and Franklin Roosevelt was not a King of England’ than to say ‘if King George VI was a President of the United States then Franklin Roosevelt was a King of England’ (notwithstanding that the speaker would take this conditional to be true, insofar as they know that both the antecedent and the consequent are false). See Chapter 4 of *Studies in the Way of Words* (1989) for Grice’s treatment of conditionals, and p. 61 for this point in particular.<sup>77</sup>

Brian Garrett has suggested that, even if Grice’s account is convincing with respect to the use of natural language conditionals in conversational contexts, it does not apply to natural language conditionals used in inner speech. Grice’s account of conditionals relies on his framework of conversational maxims, and this framework is not applicable to inner speech. Garrett actually makes this point referring to ‘beliefs’, writing that ‘beliefs are (obviously) not governed by the conversational maxims that Grice has catalogued’ (2012, p. 149). It appears, though, that he has inner speech in mind, for the example that he provides is a natural language sentence which one might ‘think to [one]self’ (2012, p. 149). In a slight error, Garrett attributes to Grice the maxim, ‘Be polite!’<sup>78</sup> He then offers this illustration of his point: ‘If, upon being introduced to a somewhat rotund individual, I think to myself, “ah, you are a somewhat rotund individual”, I violate no maxim of politeness (whereas I would if I gave voice to that thought)’ (2012, p. 149).<sup>79</sup> We can just as easily illustrate the point using one of the maxims which does appear in Grice’s framework. Suppose I am asked to name the Prime Minister of Britain and produce the inner speech utterance, ‘Sacramento is the capital of California’. I have violated no maxim of relevant, though I would have done so if I had produced the utterance aloud.

The point is not so obvious as Garrett suggests. As we have seen in this chapter, there is a detailed theory that inner speech is dialogic which, I have suggested, entails that some

---

<sup>77</sup> I have here extracted only one very small point from Grice’s discussion.

<sup>78</sup> Of course, Grice’s original framework includes no such maxim but there have been suggestions that something like a politeness maxim is necessary. See Pizziconi 2006 for an overview of the major contributions. (Grice does write the following: ‘[t]here are, of course, all sorts of other maxims (aesthetic, social, or moral in character), such as “Be polite,” that are also normally observed by participants in talk exchanges’ (1975/1989, p. 28)).

<sup>79</sup> Thanks also to Alan Hájek for discussions and his private notes on the point in the text.

episodes of inner speech are governed by Grice's framework or by some other such framework. The theory does have some motivation; this was discussed in Section 3. It has taken considerable work to show that the theory is mistaken and it has been necessary to address a number of possible objections to the argument which was made against the theory. Supposing that the point is not obvious, however, we can now see that it is a good one. It does seem that Grice's framework of conversational maxims is not applicable to inner speech. On the face of it, this means that Grice's account of conditionals cannot explain conditionals used in inner speech, which looks like a major problem for it.

## 7.2. Inner Speech and Accessibility Problems

In his (1992) paper, 'The Self as the Center of Narrative Gravity', Daniel Dennett tells what he calls elsewhere a 'Just So' evolutionary story (Dennett 1991, p. 194):

In the beginning – according to Julian Jaynes (1976), whose account I am adapting – were speakers, our ancestors, who were not really conscious. They spoke, but they just sort-of blurted things out, more or less the way bees do bee dances, or the way computers talk to each other. That is not conscious communication, surely. When these ancestors had problems, sometimes they would 'ask' for help ..., and sometimes there would be somebody around to hear them. So they got into the habit of asking for assistance and, particularly, asking questions. Whenever they could not figure out how to solve some problem they would ask a question, addressed to no one in particular, and sometimes whoever was standing around could answer them. And they also came to be designed to be provoked on many such occasions into answering questions like that – to the best of their ability – when asked.

Then one day ..., one of our ancestors asked a question in what was apparently an inappropriate circumstance: There was nobody around to be the audience. Strangely enough, he heard his own question, and this stimulated him, cooperatively, to think of an answer, and sure enough the answer came to him. He had established, without realizing what he had done, a communication link between two parts of his brain, between which there was, for some deep biological reason, an accessibility problem. One component of the mind had confronted a problem that another component could solve; if only the problem could be posed for the latter component. Thanks to his habit of asking questions, our ancestor stumbled upon a route via the ears. What a discovery! (Dennett 1992, p. 112, footnote omitted.)<sup>80</sup>

Later, Dennett continues, people realised it was not necessary to produce utterances aloud in order to achieve this effect; they could produce them silently – in inner speech. For Dennett, then, inner speech provides a solution to 'accessibility problem[s]'. It allows specialised parts of the brain to interact or communicate.<sup>81</sup>

---

<sup>80</sup> Dennett tells the same story in Dennett 1991, pp. 194-197.

<sup>81</sup> For a similar account, see Blachowicz (1997).

I do not want to make any assessment of the merits of Dennett's suggestion. However, in some respects, his suggestion resembles the theory that inner speech is dialogic. Both theories model inner speech on interpersonal communication; both make reference to the apparently dialogic or conversational structure of some series of inner speech utterances, e.g. questions followed by answers. Accordingly, one might wonder whether Dennett's suggestion is also vulnerable to the argument of this chapter.

It is not. On Fernyhough's theory, inner speech dialogue is an internalised version of interpersonal dialogue. Accordingly, inner speech dialogue depends on the same framework for regulation as interpersonal dialogue does (again, whatever exactly that framework is). This is the commitment I have focussed on in arguing against the theory in this chapter. But the theory that inner speech provides a solution to accessibility problems makes no claim implying that series of inner speech utterances which resemble interpersonal dialogue should depend on the same framework for regulation as interpersonal dialogue does. It is open to think that inner speech provides a silent 'communication link between two parts of [one's] brain' but that this 'communication link' is not regulated in anything like the way that interpersonal dialogue is. Indeed, even supposing that inner speech does provide such a 'communication link', it would require considerable additional argument to show that parts of the brain communicate in the way that distinct individuals do. Unless (improbably) one subscribes to a homunculus theory of the mind, we should probably assume that this is not the case. Accordingly, one could not argue against the theory that inner speech provides a solution to accessibility problems just by showing that the communication between brain parts which inner speech allegedly facilitates is not regulated in the same way as interpersonal dialogue.

## Chapter Three: Inner Speech and Self-Knowledge

### 1. Introduction

It is a very natural thought to suppose that inner speech might be an important source of self-knowledge, i.e. an important source of knowledge about one's own mental states. For one thing, it is certainly true that we can learn about the mental states of others by attending to the external speech that they produce, so it should not be wholly surprising if we could also learn about our own mental states by attending to the inner speech that we produce. For another, speech is especially apt for expressing beliefs, desires, fears etc., so it is tempting to think that inner speech would be well suited to a role in facilitating awareness of these when they occur in our own minds.

This natural thought, however, quickly encounters some difficulties. There is a long tradition which holds, plausibly, that our access to our own mental states is different from our access to the mental states of others. If this is so, then the appeal to the role of external speech in learning about the mental states of others will have little force. Moreover, it is certainly open to hold that, while speech is especially apt for expressing mental states once one already has knowledge of them, it plays no role in acquiring this knowledge.

For reasons that will become clear, I will refer to the view that inner speech can be a source of knowledge of our own beliefs as 'Inner Speech Interpretationism', or 'ISI'.<sup>82</sup> The purpose of this chapter is twofold: to present a problem for ISI; and to propose a supplemented version of the view which avoids the problem. I will proceed as follows. I will first sketch out the commitments of ISI drawing on the work of Gilbert Ryle, Peter Carruthers, and Quassim Cassam. Next, I will present the problem which confronts ISI. The problem – which is an adapted version of a problem arising in the epistemology of testimony – is a problem of justification. Very roughly, it notes that language can be used in misleading ways and asks, in light of this, what reason there is that anyone should believe that their inner speech can be a reliable source of self-knowledge. I will then introduce some material which will be recruited in seeking to address this problem: some material from the literature on cognitive phenomenology, and some material from John Searle's classic treatment of speech acts, especially the speech act of assertion. Drawing on this material, I will show how the phenomenology we experience when making assertions varies depending whether we are asserting propositions that we do, or do not, believe. Developing this idea further, I will

---

<sup>82</sup> I am mostly going to set aside the question of access to our other mental states and focus on beliefs. This will be explained in setting out ISI.

introduce the supplemented version of ISI – which I will call, simply, ‘Supplemented ISI’ – that avoids the problem which confronts the original version. In a word, the supplemented view is that there is something it is like to produce a sincere assertion in inner speech, and this feeling justifies one in self-ascribing belief in the proposition expressed in the assertion. I will also provide some motivation for the view. At the end of the chapter, I will offer some initial evaluation of Supplemented ISI.

It is worth saying that this chapter has a different kind of objective from Chapters One and Two. Chapters One and Two argued for particular theses about inner speech: that inner speech is a kind of actual speech; and that inner speech is not dialogic. This chapter makes no such strong claim. Rather, as the programme outlined above makes clear, it considers a problem for the view that inner speech can be a source of knowledge of our own beliefs; suggests a way the view can be amended to address that problem; and then begins to evaluate the revised view. Accordingly, this chapter might be seen as developing a view rather than as prosecuting a particular case for it.

Finally, before beginning, a point that was made towards the end of Chapter Two (in Section 5.2) should be reiterated. Our best theories of dialogue, which my argument in that chapter depended upon, seem to presuppose that we have access to our mental states independently of the speech that we produce. In this chapter, I am going to consider the possibility that inner speech might play a role in acquiring knowledge of our beliefs. Is there not a problem here? I will address this in Section 7. In short, though, the answer is that, while the theory being considered in this chapter does give a role to inner speech in acquiring knowledge of our own beliefs, it still explains why it is not possible for one to, for example, deceive oneself by producing an utterance in inner speech. Accordingly, there is no inconsistency between the point made at the end of Chapter Two and the position explored in this chapter. This also means that it is legitimate to continue using the argument strategy from Chapter Two, and I will do this at a couple of points in this chapter.

## **2. Inner Speech Interpretationism**

Before beginning any discussion of the theory that inner speech can be a source of knowledge about our own beliefs, we need to see in at least broad terms what this theory holds. The theory is most closely associated with Gilbert Ryle. In a famous passage in *The Concept of Mind*, Ryle writes:

We eavesdrop on our own voiced utterances and our own silent monologues. In noticing these we are preparing ourselves to do something new, namely to describe the frames of mind which these utterances disclose. But there is nothing intrinsically



proprietary about this activity. I can pay heed to what I overhear you saying as well as to what I overhear myself saying, though I cannot overhear your silent colloquies with yourself. (Ryle 1949, p. 184).

More recently, the theory that inner speech can be a source of self-knowledge has been developed by Peter Carruthers (2009, 2011) and Quassim Cassam (2014). Rather than summarising Ryle's, Carruthers's and Cassam's views – which would require extensive space – I am simply going to sketch a basic version of the theory that inner speech can provide access to beliefs, drawing on them from time to time.<sup>83</sup> This version will hopefully capture what is central to, and interesting about, the view, but also develop it in a way which is plausible. I call it 'Inner Speech Interpretationism', or 'ISI', for a reason that will become clear momentarily.

What does ISI hold? First, it obviously holds that inner speech can be a source of knowledge of our beliefs. But how does one get from producing and experiencing inner speech to knowledge of one's beliefs? In a nutshell, the theory holds that we do so by interpreting our inner speech. This is really the central commitment of the view – and the reason that it is named as it is.

But what exactly does the process of interpreting our inner speech involve? ISI answers this question by assimilating the process to the process of interpreting the speech of others. We get a sense of this from Ryle: 'the sorts of things that I can find out about myself are the same as the sorts of things that I can find out about other people, and the methods of finding them out are much the same' (Ryle 1949, p. 155). In a little more detail, Carruthers writes:

For present purposes, an interpretative process ... is one that accesses information about the subject's current circumstances, or the subject's current or recent behaviour, as well as any other information about the subject's current or recent mental life. For this is the sort of information that we must rely on when attributing mental states to other people. (Carruthers 2009, p. 123).

Given the importance of interpretation to ISI, it may seem evasive to characterise the process of interpreting our inner speech simply by assimilating it to another process. On a little reflection, however, it seems somewhat more reasonable. For one thing, the idea that inner speech can provide access to our own beliefs is really an application of the broader idea that we access our own beliefs in more or less the same way we access others' beliefs, including by interpreting their external speech. For another, we do not know exactly what is involved in the interpretation of the speech of others, so it would be a bold move by a proponent of ISI to advance a theory of what it is to interpret our own speech. Accordingly, ISI also holds simply

---

<sup>83</sup> There is also some overlap between the various sources. I have not sought to provide references to all sources where a particular idea occurs. Rather, I have just made use of helpful statements of the ideas in the sources.

that we interpret our own inner speech in the way that we interpret the external speech of others. Certainly, this best preserves the plausibility of the view.

Next, following both Carruthers (2011, p. 3) and Cassam (2014, pp. 138-139), ISI holds that the process of interpreting one's inner speech can be either conscious or subconscious. This relates to the last point in that the process of interpreting the external speech of others can also be either conscious or subconscious.

Before going further, something should be said about some claims that ISI does not make. ISI does not commit to a position as to whether there may be sources other than inner speech which provide access to beliefs, and it does not commit to a position as to whether inner speech may provide access to mental states other than beliefs. Most likely, a proponent of ISI would have certain views on these questions. It would be natural for a proponent of ISI to hold that our behaviour, including our external speech, as well as the other conscious mental events which take place in our minds, such as the generation of visual imagery, can provide access to our beliefs. And it would be natural to hold that inner speech can provide access to other standing mental states, such as desires, and possibly also to some episodic thoughts, such as judgements. In fact, we saw these views expressed a moment ago. Recall Carruthers: 'an interpretative process ... is one that accesses information about the subject's current circumstances, or the subject's current or recent behaviour, as well as any other information about the subject's current or recent mental life'. This obviously goes beyond inner speech. And recall Ryle: 'the sorts of things that I can find out about myself are the same as the sorts of things that I can find out about other people, and the methods of finding them out are much the same'. So, insofar as the external speech of others helps us to surmise not only their beliefs but also (for example) their desires and judgements, inner speech helps us to surmise not only our own beliefs but also (for example) our own desires and judgements. But ISI does not make these additional claims. The idea which is to be examined in this chapter is that inner speech can provide access to beliefs, and it would only reduce the plausibility of ISI if it made additional commitments.

It is helpful when thinking about self-knowledge to consider whether a particular theory makes a psychological claim (i.e. a claim about the process by which a subject attributes a mental state to herself) or an epistemological claim (i.e. a claim about how a subject can be justified in attributing a mental state to herself). ISI makes both. First, to say that we can learn about our beliefs by interpreting our inner speech is obviously to say that we can learn about our beliefs via a psychological process. But ISI also tells an epistemological story, elaborating the role that inner speech can play in justifying our self-ascriptions of beliefs. Here Cassam is helpful.

Cassam makes use of the notion of inference, which he explains as follows:

assuming that you can't know that P unless you are justified in believing that P, saying that your knowledge that P is inferential might be a way of making a point about its *justificational structure*: your knowledge that P is inferential if your justification for believing that P comes in part from having justification to believe other, supporting propositions. (Cassam 2014, p. 139; footnote omitted, emphasis added).

What are the other, supporting premises which justify us in self-ascribing a belief by making an inference from our inner speech? Cassam writes:

The inferences that lead to self-knowledge are normally mediated by the subject's implicit grasp of what is sometimes called a 'theory of mind'. You only infer from [evidence] E that you have [attitude] A because you take E to be evidence (at least in your own case) for A. Taking E to be evidence for A is an implicit theoretical commitment of yours. (Cassam 2014, p. 139; footnote omitted).

Adopting this line of thought, ISI holds the following. A subject may have a theory of mind which holds, plausibly, that subjects often express their beliefs in the speech that they produce. This is so in the case of external, interpersonal speech, in which context one is concerned to surmise the beliefs of someone else, and inner speech, in which context one is concerned to derive one's own beliefs. So, if you take it, on the basis of your theory of mind, that an utterance in inner speech provides evidence that you have a particular belief, then you are justified in ascribing this belief to yourself – at least insofar as your theory of mind is justified.

There is one final thing to say about ISI. This is that there is a point from Ryle which, adapted slightly, a proponent of ISI can adopt, but there does not seem to be a compelling reason to do so. Ryle holds that utterances of a certain kind in external speech are particularly informative about an interlocutor's mental states, i.e. those which are 'spontaneous, frank and unprepared'; likewise, he suggests that it is one's 'unstudied' utterances which are particularly informative about one's own mental states. Now, suppose someone says something to you which, you are confident, is calculated to give you a higher opinion of them than is perhaps due. Whether this remark is 'spontaneous, frank and unprepared' or 'calculated to give false impressions' is clearly something that you will be able to determine only by interpreting the remark.<sup>84</sup> A proponent of ISI will want to say that the same is true of utterances produced in inner speech: whether an utterance in inner speech really is unstudied is something that can only be determined by interpreting it. It may be that, as a matter of fact, utterances in inner speech which are unstudied are particularly informative about one's own beliefs, and it is certainly open to a proponent of ISI to say that this is so. But, on the terms of ISI, a subject does not know in advance whether an utterance is unstudied; all utterances have to be interpreted, including those which, a subject might conclude, are unstudied. So a proponent of ISI can hold

---

<sup>84</sup> Thanks to Melissa Ebbers for this point.

that unstudied utterances in inner speech are, as a matter of fact, the most informative about our beliefs, but this is not a claim which adds anything to ISI as an account of the role inner speech plays in learning about our beliefs.

In brief, then, ISI holds that inner speech can be a source of knowledge of our own beliefs. In order to provide this knowledge, inner speech needs to be interpreted. The process of interpreting one's own inner speech is the same as the process of interpreting the external speech of someone else. This process can be either conscious or subconscious. ISI does not involve any commitments as to whether sources other than inner speech can provide access to our beliefs or as to whether inner speech can provide access to mental states other than our beliefs. The theory makes both a psychological and an epistemological claim. A proponent of ISI can hold that it is one's spontaneous and frank inner speech which is most informative about one's beliefs, but this does not augment the view as an account of the role inner speech plays in learning about our beliefs.

### **3. A Problem**

Having outlined the commitments of Inner Speech Interpretationism, we can now consider how the view faces a problem. The problem is an adapted version of a problem from the epistemology of testimony, which I begin by reviewing.

#### **3.1. Vulnerability**

There is a simple problem regarding interpersonal communication which can be stated crudely as follows: why should a hearer believe what a speaker says? Jonathan Adler (2015) refers to this problem as the 'Vulnerability Problem' and gives a quite detailed statement of it:

What reason, if any, is there for a hearer to just take the speaker's word, given that the speaker is capable of lies, deception, error, and poor, ambiguous, or misleading expression? For the hearer to *trust* the speaker's word is for the hearer to ascribe *authority* to the speaker. Within the limits of presumed competence, the hearer ascribes to the speaker justification or warrant or knowledge for what she asserts. The hearer takes the speaker to be in a better position to settle the matter easily and transmit the relevant information, and so seeks the speaker's testimony (Gibbard 1990; Brandom 1994; Faulkner 2007; Keren 2007). Does the hearer have good reason to ascribe that authority? (Adler 2015, sec. 1, emphasis original).

Adler emphasises the ability of speakers to mislead hearers (intentionally or otherwise) in the first sentence of the quoted passage and the question of why a hearer should ascribe authority

to a speaker in rest of it. In this chapter, I want to address a problem which is related to the first issue Adler points to. The question why we might take anyone to have any authority about anything is not intimately related to speech, internal or external. If I follow your example in some course of action, I ascribe a certain authority to you, though not necessarily because of anything you say. The question why I take your word when you could be lying to or misleading me is a question intimately related to speech. So, focussing still on interpersonal speech, let us restate the problem focussing only on the first issue:

The Vulnerability Problem: What reason, if any, is there for a hearer to just take the speaker's word, given that the speaker is capable of lies, deception, error, and poor, ambiguous, or misleading expression?

Clearly, the Vulnerability Problem is a problem of justification. It does not suggest that we should not take a speaker's word but asks what justifies us in so doing.

The revised Vulnerability Problem applies most naturally to assertions (as, indeed, the original version does; see Adler 2015, sec. 1). It can be read as asking: what reason, if any, is there for a hearer to believe what a speaker asserts, given that the speaker is capable of lies, deception, error, and poor, ambiguous, or misleading expression? It will convenient to restrict our focus to assertions throughout most of this chapter.

Now, speakers do not ordinarily speak specifically about their mental states (although they do sometimes). Rather, they speak about the world. Nonetheless, it is natural to suppose that the utterances which speakers produce tell us something about their mental states. For example, if a speaker asserts that grass is green, it is natural to suppose that the speaker believes that grass is green. It is easy, though, to reformulate the Vulnerability Problem once more so that it applies to the vulnerability to which we are subject when we surmise a speaker's beliefs on the basis of their assertions:

The Vulnerability Problem: What reason, if any, is there for a hearer to believe that a speaker believes the proposition expressed by an assertion that she makes, given that the speaker is capable of lies, deception, error, and poor, ambiguous, or misleading expression?

Future references to the Vulnerability Problem are to this version of it.

### **3.2. Vulnerability and Inner Speech**

The central claim of ISI, of course, is that inner speech can be a source of knowledge about one's own mental states. We can now see how the difficulty for ISI starts to emerge. The

notion that a subject can acquire knowledge of her own mental states by interpreting her inner speech is subject to something very similar to the Vulnerability Problem, notwithstanding that the speaker and the hearer are the same person. We can call the problem for ISI the ‘Vulnerability (IS) Problem’, the ‘IS’ signifying inner speech. We can set it out provisionally as follows, though we will see momentarily that it requires a refinement:

The Vulnerability (IS) Problem: What reason, if any, is there for a subject to believe that she believes the proposition expressed by an assertion that she makes in inner speech, given that the subject is capable of lies, deception, error, and poor, ambiguous, or misleading expression?

Now, one might take pause here. It was emphasised a number of times in Chapter Two that one cannot successfully lie to oneself in inner speech: there is no one who is liable to be deceived in inner speech, as there is in interpersonal conversation. The point was made to illustrate the idea that the framework of shared assumptions which makes interpersonal conversation possible (whatever exactly this framework is) simply has no application in the context of inner speech. Yet now the statement of the Vulnerability (IS) Problem seems to take it for granted that one can lie and deceive etc. in inner speech. Is this not a contradiction?

Of course, we can produce assertions in inner speech which express our beliefs, and we can produce assertions in inner speech which do not express our beliefs. The point urged in Chapter Two is that no one will ever be misled when they produce an assertion in inner speech which does not express a belief that they hold. It is for this reason that one can never lie successfully in inner speech. The Vulnerability (IS) Problem, by contrast, highlights a point which applies to those assertions which we produce in inner speech and which *do* express our beliefs. The point is that, when a subject produces an assertion in inner speech which does express a belief they hold, it seems that the subject somehow knows this, even though they are also capable of producing assertions in inner speech which do not express their beliefs. The Vulnerability (IS) Problem asks how this can be so; it seeks some justification for the subject’s knowledge.

The problem should be restated in a way which avoids the potential confusion, and we will see momentarily that the amendment is actually important for another reason. The problem can now be stated as follows:

The Vulnerability (IS) Problem: What reason, if any, is there for a subject to believe that she believes the proposition expressed by an assertion that she makes in inner speech, in those circumstances in which she does believe the proposition expressed by the assertion, given that she is capable of making assertions in inner speech which express propositions that she does not believe?

The Vulnerability (IS) Problem, like the Vulnerability Problem, is a problem of justification. It does not suggest that we should not self-ascribe beliefs on the basis of assertions we make in inner speech; rather, it asks what justifies us in doing so.

Of course, ISI does have something to say about justification. It claims that, if you take it, on the basis of your theory of mind, that an utterance in inner speech is evidence that you have a particular belief, then you are justified in ascribing this belief to yourself – at least insofar as your theory of mind is justified. Before going further, it is important to note that this account does not solve the Vulnerability (IS) Problem. This is because of the amendment we have just made to the Vulnerability (IS) Problem. A general theory of mind might provide a reason for ascribing to others belief in the propositions expressed in assertions that they make. But the (newly amended) Vulnerability (IS) Problem asks for justification for self-ascribing beliefs in propositions expressed by assertions made in inner speech, but only *in those circumstances in which we do believe the proposition expressed by the assertion*. If someone else asserts in external speech, ‘Grass is blue’, knowing that this is false, we are still justified in ascribing to them the belief, ‘grass is blue’, because our theory or mind holds that other people generally believe what they assert. But the situation is different with inner speech. An individual who makes the assertion, ‘Grass is blue’, in inner speech, knowing that this is false, is not justified in self-ascribing belief in the proposition, grass is blue. Yet an individual who makes the assertion, ‘Grass is green’, and who believes this, is somehow justified in self-ascribing that belief. The Vulnerability (IS) Problem asks that this justification be explained – and appealing to a general theory of mind is not sufficient to do this.

In the discussion that follows, I will tend to drop the somewhat laboured language of the formal statement of the Vulnerability (IS) Problem and simply speak, for example, of why it is that one should believe that one believes one’s assertions in inner speech. It should be clear, though, that locutions such as this are just a shorthand for the formal statement of the Vulnerability (IS) Problem.

### **3.3. Borrowing a Solution?**

There has been considerable philosophical discussion as to what justifies us in taking others at their word – i.e., roughly, as to what solves what I am calling the Vulnerability Problem. Is it possible that a solution that has been proposed in this context might be available for addressing the Vulnerability (IS) Problem? In this subsection, I am going to demonstrate that there is a particular difficulty that would emerge if we sought to borrow a *certain kind* of proposed justification for taking others at their word, and a more general difficulty that would

emerge if we sought to make use *any kind* of proposed justification for taking others at their word.

### 3.3.1. Ross and Rules of Language

First, any proposed justification for taking others at their word which emphasises the use of language as a means of interpersonal communication will be unavailable in addressing the Vulnerability (IS) Problem. An elegant example of such an account is offered by Angus Ross (1986).

Ross observes that we cannot legitimately criticise a naturally occurring sign which somehow causes us to be misled. He provides an example:

Suppose we find a natural sign occurring in the absence of the phenomenon of which it is a sign, say (what looks like) smoke without fire. We may not be pleased to discover we have been misled, if we have been misled, but *criticism*, at least of the smoke or of what is producing it, will be out of place. (Ross 1986, p. 77; emphasis original).

By contrast, Ross notes, someone who speaks falsely can legitimately be criticised, whether they have done so intentionally or otherwise. Why is this so? The reason is that language use, as a communal practice, requires that we generally speak truthfully: '[t]he point is that without an agreed distinction between correct and incorrect ways of using language it would not be a possible means of communication' (1986, p. 77). Speaking falsely is an 'incorrect use of language' because language would not be a possible means of communication if people did not generally speak truthfully (1986, p. 77). (Ross's claim here is presumably not meant to apply to all uses of language. No one should think that a writer of fiction uses language incorrectly.<sup>85</sup> The idea is simply that linguistic communication is possible only if we ordinarily speak truthfully.)

Ross explains his point:

One way of putting this is to say that the regularities on which the hearer relies are more than *de facto* regularities. Speech is a rule-governed activity and its rules impose certain normative requirements on speakers. (1986, p. 77).

Now, if speech is a rule-governed activity, this has implications for the relationship between speaker and hearer. Ross writes:

It is a quite general feature of rule-governed life that the responsibility for ensuring that one's actions conform to the rules lies primarily with oneself and that others are

---

<sup>85</sup> Thanks to Daniel Stoljar for this point.



in consequence entitled to assume, in the absence of definite reasons for supposing otherwise, that one's actions do so conform. (1986, p. 77).

This makes sense. In any rule-governed activity, one can hardly be responsible for the conformity of someone else's actions to the relevant rules. The rules of language require that we speak truthfully; if they did not, language would not be a 'possible means of communication'. It is the speaker's responsibility for speaking truthfully which justifies a hearer in believing what the speaker says. 'The speaker', Ross writes, 'in taking responsibility for the truth of what he is saying, is offering his hearer ... a *guarantee* that it is true' (1986, pp. 79-80; emphasis original). We can add that the guarantee also provides a reason for a hearer to believe that the speaker believes the assertion that they make. For how could a speaker issue a guarantee that something is true without thereby also warranting that they believe it to be true? So, Ross's account is a plausible solution to the Vulnerability Problem.

It is an attractive account but it cannot be extended to solve the Vulnerability (IS) Problem. The reason is that it depends critically on the way that language is used in a *community*. Issuing a guarantee for the truth of what one says – and, by extension, a guarantee that one believes what one says – is something that one can only do *vis-à-vis* another person. There is no one who would have a claim against a speaker for shirking their responsibility, or for defaulting on a guarantee they had issued, if they said something in inner speech that they did not believe. In consequence, one cannot rely on a guarantee one has issued in order to conclude that one believes something they have asserted in inner speech. (This resembles, of course, the main argument strategy in Chapter Two. Again, I will return to the relationship between that chapter and this one in Section 7).

This difficulty will arise in attempting to address the Vulnerability (IS) Problem by recruiting any proposed solution to the question of what justifies us in taking others at their word which depends on the social aspect of interpersonal speech. Inner speech simply does not have this social aspect.

### **3.3.1. Burge and the Acceptance Principle**

Are there any proposed solutions to the question of what justifies us in taking others at their word which do not depend on the social aspect of interpersonal speech? And might they possibly be recruited in solving the Vulnerability (IS) Problem? This approach does look somewhat more promising.

Consider Tyler Burge's (1993) treatment of testimony. Burge writes that '[j]ustification in acquiring beliefs from others may be glossed, to a first approximation', by what he calls the 'Acceptance Principle':

A person is entitled to accept as true something that is presented as true and that is intelligible to him, unless there are stronger reasons not to do so. (Burge 1993, p. 467).

Burge emphasises that the Acceptance Principle is 'not an empirical principle' (1993, p. 469). It is not justified by the relative frequency with which one finds that the assertions of others are true. Rather, 'the general form of justification associated with the principle' is as follows:

A person is apriori entitled to accept a proposition that is presented as true and that is intelligible to him, unless there are stronger reasons not to do so, because it is prima facie preserved (received) from a rational source, or resource for reason; reliance on rational sources – or resources for reason – is, other things equal, necessary to the function of reason. (1993, p. 469).

Burge differentiates rational sources and resources for reason. Resources for reason, such as perception and memory, 'need not themselves be rational', but 'may provide material and services that a rational being is apriori entitled to rely upon' (1993, p. 470). 'Rational sources', by contrast, 'are sources that themselves are a capacity to reason or are rational beings' – which would presumably include most, or maybe all, speakers with whom one might converse (1993, p. 470). It is testimony from rational sources which is of present concern.

Burge's argument that the Acceptance Principle is justified proceeds as follows. First, he writes, 'if something is a rational source, it is a prima facie source of truth. For a condition on reasons, rationality, and reason is that they be guides to truth' (1993, p. 470). Accordingly, if you take someone to be a rational source, then you should generally take them to be speaking truthfully, absent some reason not to. '[I]t is rational', he writes, 'to accept the affirmed deliverances of a rational source', absent some reason not to accept them (1993, p. 471).

Burge adds that

[i]t is not just the rationality of a source that marks an apriori prima facie connection to truth. The very content of an intelligible message presented as true does so as well. For content is constitutively dependent, in the first instance, on patterned connections to a subject matter, connections that insure in normal circumstances a baseline of true thought presentations. (1993, p. 471).

Content would not be possible if we could not generally assume of intelligible messages presented as true – or, we might say, of intelligible assertions – that they are true. For if we could not generally make such an assumption of intelligible messages, then there would be no reason to treat them as anything other than series of sounds (or print on a page etc.).

Finally, one can prima facie assume that someone who intelligibly presents something as true is a rational source. For the ability to intelligibly present something as true

presupposes at least a derivative connection to a system of perceptual, cognitive, and practical interactions with a world, involving beliefs and intentional activity. Belief and intention in turn presuppose operation under norms of reason or rationality – norms governing information acquisition, inference, and practical activity. (1993, pp. 471-472).

Put a little more simply: if a speaker can present some content intelligibly, the speaker must have (among other things) beliefs and intentions; otherwise, how could they do so? But having beliefs and intentions requires operating according to norms which govern the acquisition of information and the drawing of inferences etc. – i.e. to norms of rationality.

In brief, then, a rational source is a *prima facie* source of truth, because being a rational source means being governed by norms for truth-seeking. Also, messages presenting intelligible content as true can, *prima facie*, be assumed to be true. If they could not, then we would have no reason not to treat such messages as, for example, mere series of sounds. Finally, one can presume that a speaker who presents content intelligibly is rational, because presenting content intelligibly is only possible if one has beliefs and intentions – mental states which one can have only if one acquires information, draws inferences etc. in a way which is governed by norms of rationality.

As in the case of Ross's account, a small extension can be made so that the account provides a plausible solution to the Vulnerability Problem. If one has a *prima facie* entitlement to believe a proposition because it was presented as true by a rational source, then one must also have a *prima facie* entitlement to believe that the source believes the proposition. If you doubt that a particular source believes a proposition which they present as true, then you could hardly be justified in accepting the proposition as true, on the basis that it was presented as true by that source.

Burge's is another elegant account. Also, we can see that it depends not on the social aspect of interpersonal speech but on the rationality of individuals. Actually, Burge makes it explicit that the account does not rely on the social aspect of inner speech. He writes:

The Acceptance Principle and its justification are formulated so as to be neutral on whether what is 'presented as true' comes from another person. Its application does not depend on an assumption that the source is outside oneself (although further articulation will, I think, give this source a place in the account). Many of the differences between content passing between minds and content processed by a single mind derive from differences in modes of acquisition and in necessary background conditions, that do not enter into the justificational force underwriting an entitlement. (1993, p. 474).

And, shortly afterwards:

One can presume that a presentation of something as true by a rational being – whether in oneself or by another – has, *prima facie*, something rationally to be said for it. (1993, p. 475).

One might take pause because of the parenthetical remark from the first quotation in the second half of this sentence: '[The Acceptance Principle's] application does not depend on an assumption that the source is outside oneself (although further articulation will, I think, give this source a place in the account)'. It is not apparent from the text what Burge is referring to here; he does not elaborate on it in 'Content Preservation'. In any event, it seems entirely clear that Burge takes the Acceptance Principle to be applicable even when the source is oneself.

I think that Burge's Acceptance Principle actually cannot be recruited in addressing the Vulnerability (IS) Problem – that it is not applicable when the source is oneself – and the reason relates to the final amendment that was made to the formal statement of that problem. This issue was discussed in some detail above with respect to the idea that a general theory of mind might justify self-ascriptions of belief on the basis of assertions in inner speech, but it will be worthwhile working through it again. What do we want when we ask what justifies us in taking others at their word? We want a general reason to believe that a speaker speaks truthfully and that they believe the propositions expressed by assertions that they make. The situation is different with inner speech. We do not want a general reason to believe that what we assert in inner speech is true. This would justify too much. If we had such a reason, then one could assert in inner speech, 'Grass is blue', and have a reason to believe that this is true. But nor do we want a general reason to believe that we believe what we assert in inner speech. This would still justify too much. If we had this kind of reason, then one could assert in inner speech, 'Grass is blue', and have a reason to believe that one believes this – even if one does not in fact believe it. With respect to inner speech, what we really want is a justification to believe that we believe what we assert in inner speech, but only in the circumstances in which we do actually believe what we assert. This is what the Vulnerability (IS) Problem asks for.

So, it may be the case that Burge's account explains why we are entitled to believe what others say, and to believe that they believe what they say. However, the account cannot be recruited in solving the Vulnerability (IS) Problem. For, applied to inner speech, it would provide the wrong kind of justification. It might provide a reason to believe that we believe what we assert in inner speech, but it would not provide this reason only in the circumstances in which we do actually believe what we assert in inner speech. One does not do any better by appealing to Burge's Acceptance Principle than one does by appealing to a general theory of mind.

Finally, notice that this is a general feature of any account which promises to explain why we are justified in taking others at their word. Such an account may provide a justification for generally believing that another speaker believes what they say, but not only in the circumstances in which that other speaker does actually believe what they say. And this is what is needed if such an account is to be adapted to solve the Vulnerability (IS) Problem.

So, the situation is this. There is a specific reason that any proposed solution to the question why we are justified in taking others at their word which depends critically on the social aspect of interpersonal speech cannot be recruited in addressing the Vulnerability (IS) Problem: inner speech simply does not have this social aspect. Further, there is a general reason that *any* proposed solution to the question why we are justified in taking others at their word cannot be recruited to solve the Vulnerability (IS) Problem: it will not be able to explain why we are justified in believing that we believe what we assert in inner speech, with the appropriate restriction that we are so justified only when we do in fact believe what we assert. We will have to set aside any thought of borrowing a solution, and search for a solution which is specific to the Vulnerability (IS) Problem.

#### **4. A Detour**

Shortly, I will begin to develop a supplemented version of ISI which has some defence against the Vulnerability (IS) Problem. I will call this new version of ISI ‘Supplemented ISI’. As preparation, it is necessary to take a detour through a couple of other areas of philosophy in order to introduce some material which will play important roles in Supplemented ISI. This material comes from two sources: recent work on cognitive phenomenology, and foundational work on speech acts.

##### **4.1. Cognitive Phenomenology**

The idea that there is cognitive phenomenology – that there is something distinctly cognitive that it is like to have an occurrent, conscious thought – has received considerable attention in recent philosophy of mind (for a collection on the topic, see Bayne & Montague 2011a). David Pitt (2004, 2011) has made some ambitious claims about cognitive phenomenology. The one which is relevant for present purposes is that it is the cognitive phenomenology of conscious thoughts which allows us to know what those thoughts are (Pitt, 2004, 2011).

Putting things very roughly, Pitt believes that it is ‘obvious’ that all conscious mental states have phenomenal properties, including conscious thoughts (Pitt 2004, p. 3), where conscious thoughts are mental states with propositional content (2004, p. 4). And, certainly, it is true that at least very many types of non-cognitive conscious mental states have phenomenal properties, e.g. ‘sensations, perceptions, proprioceptions, emotions’ (2004, p. 3). But this is not all. For Pitt, the phenomenology of a conscious thought allows one to know a great deal about that conscious thought. Just as it is the phenomenology of, say, hearing a dog bark that allows one

to know that one is hearing a dog bark, it is the phenomenology of judging that coffee is bitter that allows one to know that one is judging that coffee is bitter (Pitt 2004, 2011).

Pitt's theory is not, strictly speaking, in competition with ISI. Pitt's theory is about access to occurrent conscious thoughts; ISI is a theory about access to standing beliefs. In any event, for present purposes, I just want to review one argument strategy which is relied on by those who believe in cognitive phenomenology. The strategy involves contrasting examples of experiences which seem to vary in phenomenology, even though they have exactly the same sensory content (Bayne & Montague 2011b, p. 22).

Consider first a situation in which two people hear a sentence spoken in a language which only one of them knows. One will understand what has been said and the other will not. But it seems there will be a further difference: their experiences will *feel* different. Moreover, insofar as the two have had exactly the same sensory experience, the difference in phenomenology cannot be sensory. Rather, the difference, it seems, must be in the quality of their cognitive experiences: the experience of understanding in one case and of not understanding in the other (Bayne & Montague 2011b, p. 22, though the example has been used previously; Bayne & Montague 2011b cite Moore 1953, Strawson 1994, and Siewert 1998 as examples).

A related example emphasises the difference in an individual's experience before and after 'the sudden and acute onset of an experience of grasping something' (Kriegel 2011, p. 92). One can generate the relevant contrast by reading this passage:

A newspaper is better than a magazine. A seashore is a better place than the street. At first it is better to run than to walk. You may have to try several times. It takes some skill but it is easy to learn. Even young children can enjoy it. Once successful, complications are minimal. Birds seldom get too close. Rain, however, soaks in very fast. Too many people doing the same thing can also cause problems. One needs lots of room. If there are no complications it can be very peaceful. A rock will serve as an anchor. If things break loose from it, however, you will not get a second chance. (Kriegel 2011, p. 92, quotes this passage citing Mangan 2001 and Chudnoff 2013).<sup>86</sup>

On a first reading, the passage is baffling. If you are told that the passage is about kites, however, and then reread the passage, the experience will feel very different. Again, though, it is not the sensory experience which feels different: it is the same in both instances. The difference is that you have an experience of grasping the meaning of the passage on the second

---

<sup>86</sup> Though it appears that the series of sentences was first utilised in a study out by Bransford & Johnson 1972.

Chudnoff's paper was forthcoming at the time Kriegel's was published, which is why it postdates Kriegel's.

reading – an experience which is different in character from the experience of not grasping the meaning of the passage.

Another nice example contrasts the experience of getting a joke with the experience of not getting the same joke (Smithies 2013, p. 744; Siewert 2011; Horgan & Potrč 2010). Take the following joke (not, I regret, my own):

There's a band called 1023MB. They haven't had any gigs yet.

To get the joke, one needs to realise that 'gig' is a colloquial term for 'gigabyte' as well as a word which means 'musical performance'; that 'MB' is an abbreviation for 'megabyte'; and that a megabyte consists of 1024 gigabytes. The sensory experience of hearing the words spoken to you will be the same whether or not you get the joke but, again, the experience will feel different in the two situations.

One point should be noted before moving on. Some philosophers believe that the contrast in phenomenology in pairs of cases like those reviewed can be accounted for without appealing to cognitive phenomenology. Consider again the case of the two people who hear a sentence spoken in a language which only one of them knows. Tye & Wright (2011) suggest that the difference in the experiences of the two individuals is actually just a difference in sensory phenomenology. They write:

When we hear someone speaking in a language we understand, the phonological processing of the sound stream is different from the processing that goes on when we hear someone speaking in a language we do not comprehend. ... The result is that the auditory experiences we undergo [when we hear someone speaking in a language we understand] are different from those we would undergo were we to hear the same sound stream without understanding the language. (Tye & Wright 2011, p. 337).

For the purposes of this chapter, it is actually not necessary to settle the question whether the difference between the experiences of the individuals in this example – and between the first and second experiences of the individuals in the other pairs of cases – is a difference in cognitive phenomenology, or whether it can be otherwise accounted for. All that is necessary is that there is some difference in phenomenology – and this much seems clear. Accordingly, one should not necessarily find the use of this argument strategy in what follows objectionable, even if one happens not to believe in cognitive phenomenology.

## **4.2. Speech Acts**

My treatment of speech act theory will be similarly brief. Again, I simply want to review some (this time, very foundational) material which will be important in the next chapter. The classic

work on speech acts is John Searle's (1969) *Speech Acts: An Essay in the Philosophy of Language*, and it will suffice to extract a few points from this treatment of the topic.

Searle's book is (of course) a study of speech acts. What he means by 'speech acts' (or 'linguistic acts') is the acts that we typically perform when we utter words and sentences: not only producing sounds (or written words etc.) and referring to and describing objects, but also acts such as making assertions, asking questions, making promises etc. And '[t]he reason for concentrating on ... speech acts', Searle writes, 'is simply this: all linguistic communication involves linguistic acts' (1969, p. 16). For Searle, then, understanding speech acts is crucial to understanding language. Searle acknowledges that language has 'formal features which admit of independent study' and that '[a] great deal can be said in the study of language without studying speech acts' (1969, p. 17). But he writes that:

a study purely of those formal features, without a study of their role in speech acts, would be like a formal study of the currency and credit systems of economies without a study of the role of currency and credit in economic transactions. (1969, p.17).

Such a study would be 'necessarily incomplete' (1969, p.17).

On Searle's account, when individuals do make statements, ask questions, issue commands, make promises etc., they perform at least three types of act, which he classifies as follows:

- (a) Uttering words (morphemes, sentences) = performing *utterance acts*.
- (b) Referring and predicating = performing *propositional acts*.
- (c) Stating, questioning, commanding, promising, etc. = performing *illocutionary acts*. (Searle 1969, p. 24; emphasis original).

Searle is quick to ward off a possible misunderstanding of the relationship between these three kinds of acts:

I am not saying, of course, that these are separate things that speakers do, as it happens, simultaneously, as one might smoke, read and scratch one's head simultaneously, but rather that in performing an illocutionary act one characteristically performs propositional acts and utterance acts. (1969, p. 24).

Searle completes his taxonomy with 'perlocutionary acts',<sup>87</sup> which he introduces as follows:

Correlated with the notion of illocutionary acts is the notion of the consequences or *effects* such acts have on the actions, thoughts, or beliefs, etc. of hearers. For example, by arguing I may *persuade* or *convince* someone, by warning him I may *scare* or *alarm* him, by making a request I may *get him to do something*, by informing him I may *convince him* (*enlighten, edify, inspire him, get him to realize*). (1969, p. 25; emphases original).

---

<sup>87</sup> Searle adopts the notions of the illocutionary act and the perlocutionary act (as well as these expressions) from Austin (1962) (see Searle 1969, pp. 23, 25, though cf. Footnote 1 on p. 23 of Searle's text).



It is important that one can perform different illocutionary acts by performing the same propositional act. Take the following utterances produced in appropriate circumstances:

1. Sam smokes habitually.
2. Does Sam smoke habitually?
3. Sam, smoke habitually!
4. Would that Sam smoked habitually. (1969, p.22).

Producing each of these utterances would amount to performing the same propositional act: the same object (Sam) is referred to and the same predicate is expressed about that object (that he smokes). Clearly, though, 1. is an assertion or statement; 2. is a question; 3. is an imperative or command; and 4. is a wish.

In a similar vein, performing the same utterance act can amount to performing different propositional acts and illocutionary acts (1969, pp. 24-25). Take the sentence, 'You need to settle on the land'. It consists of the same words. However, it could express either of two propositions, because 'settle' is ambiguous. It could express the proposition that you inhabit a piece of land, or the proposition that you finalise the sale of a piece of real estate. And, in each case, the proposition might be asserted or commanded.

What is especially relevant for present purposes is Searle's treatment of a specific illocutionary act, assertion. Searle sets out the necessary and sufficient conditions for the performance of a number of illocutionary acts. On his account, a speaker, *S*, makes an assertion to a hearer, *H*, if and only if the speaker expresses '[a]ny proposition *p*' (1969, p. 66) – i.e. performs a propositional act expressing *p*, where this involves referring to an object and expressing some predicate of it – and:

1. (a) '*S* has evidence (reasons etc.) for the truth of *p*'; and (b) '[i]t is not obvious to both *S* and *H* that *H* knows (does not need to be reminded of, etc.) *p*' (1969, p. 66);
2. '*S* believes *p*' (1969, p. 66); and
3. *S*'s propositional act expressing *p* '[c]ounts as an undertaking to the effect that *p* represents an actual state of affairs' (1969, p. 66).

Condition 1. is called the '[p]reparatory' condition; 2. is called the '[s]incerity' condition; and 3. is called the '[e]ssential' condition (1969, p. 66).

There are utterances which do not satisfy all of the conditions just iterated and which we would naturally think of as assertions. Keeping with the classroom theme, suppose a student is asked by her teacher, 'How many elements are there on the periodic table?'. She answers, 'There are 118'. The student knows that her teacher knows how many elements there are on the periodic table, so condition 1(b) is not satisfied.<sup>88</sup> I am going to set this point aside; we will see in the next section that it loses relevance in the context of inner speech. A more important point for

---

<sup>88</sup> Thanks to Daniel Stoljar for this point. See Searle 1969, p. 69 for a related point and similar example.

present purposes is that the effect of the second condition, the sincerity condition, is that there cannot be assertions which a speaker does not believe. Suppose a child says to her mother, ‘I didn’t take the cookie’, fully aware that she did take the cookie. On Searle’s account, this would not be an assertion.<sup>89</sup> Here I will depart from Searle, allowing that there can be sincere and insincere assertions, depending whether or not the sincerity condition is satisfied.

## 5. The Phenomenology of Speech

Adopting the strategy, reviewed in Section 4.1, which is sometimes used to in an effort to demonstrate the existence of cognitive phenomenology, I want to use a pair of cases to show how the phenomenology of making an assertion can vary depending whether or not all of the conditions for performing that kind of illocutionary act are satisfied. This will complete the preparation for showing, in the next section, how ISI can be supplemented so that it can offer an answer to the Vulnerability (IS) Problem.

Consider the following pair of cases:

1. Elizabeth is a solicitor talking about some co-workers with another colleague, Galina. Elizabeth knows that Galina is friends with the human resources manager, Eric. At one point, Galina says, ‘Eric is such a nice guy’. Elizabeth replies, ‘He’s really smart too’. Elizabeth believes that Eric is really smart.
2. Elizabeth is a solicitor talking about some co-workers with another colleague, Galina. Elizabeth knows that Galina is friends with the human resources manager, Eric. At one point, Galina says, ‘Eric is such a nice guy’. Elizabeth replies, ‘He’s really smart too’. Elizabeth does *not* believe that Eric is really smart.

In Case 1, the sincerity condition for Elizabeth’s assertion is satisfied (it is a sincere assertion); in Case 2, the sincerity condition is not satisfied (it is an insincere assertion). What I wish to suggest now is that the phenomenology that Elizabeth would experience in the two cases would be different. Producing the assertion in which the sincerity condition is not satisfied would feel different from producing the assertion in which it is. I expect this point will be obvious on reflection.

Is the difference in phenomenology a difference in cognitive phenomenology? How one sees this is likely to depend on one’s background commitments. Believers in cognitive phenomenology may be inclined to think that the contrast between the cases provides more evidence for their claims. They will say that there would be a difference in Elizabeth’s

---

<sup>89</sup> Again, thanks to Daniel Stoljar for this point.

phenomenology in Cases 1 and 2 but it is not a difference in sensory phenomenology. Elizabeth produces the same words in the two cases, so she would have the same kinaesthetic experiences (of speaking) and the same auditory experiences (of hearing her speech). Deniers of cognitive phenomenology might seek to identify some non-obvious difference in the Elizabeth's sensory experiences in Case 1 and Case 2, in the style of Tye & Wright's treatment of the phenomenology of understanding reviewed above. In any event, all that actually matters for present purposes is that there would be a difference in the quality of Elizabeth's experience in Cases 1 and 2. This much seems to be true.

What about inner speech? It may be that Searle's style of analysis of assertions does not transfer very well to inner speech, for it clearly presupposes that the speaker is a different individual from the hearer. There are different possible explanations for this. One might think that, in the absence of distinct speakers and hearers, utterances produced in inner speech simply cannot be speech acts. This would do no damage to Searle's project: recall that his reason for focussing on speech acts was that 'all linguistic *communication* involves linguistic acts' (Searle 1969, p. 16; emphasis added.) Alternatively, one might think that the failure to accommodate utterances produced privately – whether externally or internally – is a shortcoming of speech act theory.

We can set aside the question whether (and, if so, how well) Searle's analysis of the speech act of assertion, presupposing as it does that a distinct speaker and hearer, transfers to inner speech. Rather, let us see what Searle's analysis of assertions would look like if we remove this assumption. In practical terms, this will simply mean removing the second part of the preparatory condition, i.e. that '[i]t is not obvious to both *S* and *H* that *H* knows (does not need to be reminded of, etc.) *p*'. (This is the reason foreshadowed above that the second part of the preparatory condition loses relevance in the context of inner speech.) The analysis of assertions would be as follows: a speaker, *S*, makes an assertion if and only if the speaker expresses '[a]ny proposition *p*' (Searle 1969, p. 66) – i.e. performs a propositional act expressing *p*, where this involves referring to an object and expressing some predicate of it – and:

1. '*S* has evidence (reasons etc.) for the truth of *p*' (1969, p. 66);
2. '*S* believes *p*' (p. 66); and
3. *S*'s propositional act expressing *p* '[c]ounts as an undertaking to the effect that *p* represents an actual state of affairs' (p. 66).

Let us still allow that there can be sincere and insincere assertions in inner speech, depending whether or not the sincerity condition is satisfied.

Plausibly, the references to ‘an undertaking’ in the revised analysis still suggests that an assertion requires a distinct speaker and hearer. If so, then it might be preferable to replace ‘an undertaking to the effect’ with ‘a statement to the effect’ or ‘a remark to the effect that’. But one can substitute whichever locution one likes.

I now want to focus on another pair of cases. Consider the following:

3. Lena is planning a holiday in Germany. She wants to visit Munich, Düsseldorf, and Cologne, but is uncertain about visiting Berlin, where she has spent time previously. She produces the inner speech utterance, ‘I really liked Berlin when I was there before’. Lena did really like Berlin when she was there before.
4. Lena is planning a holiday in Germany. She wants to visit Munich, Düsseldorf, and Cologne, but is uncertain about visiting Berlin, where she has spent time previously. She produces the inner speech utterance, ‘I really liked Berlin when I was there before’. Lena did *not* really like Berlin when she was there before.

Clearly, the sincerity condition from the analysis of assertions, now slightly revised to apply in the context of inner speech, is satisfied in Case 3 but not in Case 4. Again, I suggest, we can expect that there will be a difference in the quality of Lena’s experience in the two cases, more or less parallel to the difference in the quality of Elizabeth’s experience in Cases 1 and 2.

## **6. Supplemented ISI**

It is now possible, finally, to introduce what I call ‘Supplemented ISI’, a version of ISI which can offer an answer to the Vulnerability (IS) Problem. I will do this in the first subsection. In the second, I will provide some initial motivation for it. I will then briefly suggest that Supplemented ISI likely has broader application than might immediately be apparent.

### **6.1. Introducing Supplemented ISI**

After all of the preparatory work above, it is possible to suggest a very simple supplementation to ISI in order to deal with the Vulnerability (IS) Problem. Recall once more that problem:

The Vulnerability (IS) Problem: What reason, if any, is there for a subject to believe that she believes the proposition expressed by an assertion that she makes in inner speech, in those circumstances in which she does believe the proposition expressed

by the assertion, given that she is capable of making assertions in inner speech which express propositions that she does not believe?

The supplementation to ISI which I want to suggest is as follows. There is something it feels like to produce an assertion in inner speech which expresses a proposition you believe. If you have that feeling when you produce an assertion in inner speech, then you are justified in believing that you believe the proposition expressed by that assertion, i.e. in ascribing that belief to yourself. This is so notwithstanding that you are capable of making assertions in inner speech which express propositions that you do not believe. ISI, so supplemented, is what I call ‘Supplemented ISI’.

It is important to note that it is not just the phenomenology of producing a sincere assertion which justifies self-ascribing belief in the proposition expressed.<sup>90</sup> The content of the assertion matters. The reason is that the phenomenology of producing a sincere assertion in inner speech remains constant across different assertions (except, of course, for the sensory phenomenology). You can produce any sincere assertion and the phenomenology of producing a sincere assertion will remain the same.

Appealing to a kind of phenomenology which is often taken to be cognitive would not be a natural manoeuvre for a proponent of ISI. As we noted in setting out ISI, those who attribute a role to inner speech in the acquisition of self-knowledge might also afford a role to one’s own behaviour, which can be observed through perceptual channels, or to other forms of sensory imagery – although these are not actually tenets of ISI itself. Nothing about ISI suggests a role for cognitive phenomenology, which really has nothing in common with these other possible sources of self-knowledge. There are two things that could be said here by a proponent of ISI who sees things this way.

First, they could insist that the phenomenology we experience when making an assertion in inner speech which satisfies all of the relevant conditions, as opposed to that which we experience when making an assertion in inner speech which does not, is actually a difference in sensory phenomenology. As foreshadowed, all that matters is that there is something phenomenologically distinctive about producing in inner speech an assertion which one believes. This approach is continuous with the approach of Tye & Wright (2011) in their treatment of the difference in phenomenology two people would experience upon hearing an utterance in a language which only one of them understands. In order to proceed in this way, however, one would need to provide some plausible account as to how there is a difference in the sensory phenomenology we experience when we produce assertions in inner speech which

---

<sup>90</sup> Thanks to Daniel Stoljar for bringing this concern to my attention.

either do, or do not, satisfy all the relevant conditions. It is far from obvious what type of account could be offered.

Alternatively, they could just accept that the relevant phenomenology is cognitive and that this cognitive phenomenology does play the suggested role in acquiring knowledge of our own beliefs. I suggest that this would involve making only a minimal concession. It does not entail, for example, that occurrent mental states each have a unique phenomenology by which we can identify them. It just requires accepting that one kind of assertion made in inner speech is reliably accompanied by a particular kind of cognitive phenomenology, and another kind of assertion is reliably accompanied by a different kind of cognitive phenomenology. And the contrast cases considered above surely do provide a quite strong, albeit not a conclusive, reason to believe that this is so.

## 6.2. Phenomenology and Justification

At this point, I have sketched out Supplemented ISI, which was the primary objective of this chapter. I now want to offer one line of argument supporting it, in order to at least establish its plausibility. In a word, this is to show that Supplemented ISI provides a convincing account of why an example of what is sometimes called a ‘Moore-paradoxical sentence’ seems irrational.

Gareth Evans famously wrote the following:

in making a self-ascription of belief, one’s eyes are, so to speak, or occasionally literally, directed outward – upon the world. If someone asks me ‘Do you think there is going to be a third world war?’, I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question, ‘Will there be a third world war?’ I get myself in a position to answer the question whether I believe that  $p$  by putting into operation whatever procedure I have for answering the question whether  $p$ . (Evans 1982, p. 225, as cited in Silins 2012, p. 299).

This is the classic statement of a so-called ‘transparency’ view of self-knowledge. What it offers, though, is an account as to how one might come to know that they have a particular belief, not an account as to how one might be *justified* in self-ascribing that belief. These are different questions. (Consider: Someone might be able to explain how they came to believe that there are unicorns, but it does not follow that they are justified in believing that there are unicorns.) It is the question of justification in self-ascriptions of belief which is of current interest.

Nico Silins (2012) advances an account derived from Evans’s which is concerned with the justification of self-ascriptions of belief. He states his position carefully and precisely but I am going to leave aside the details, because I think that there is a problem with his central

idea. Put broadly, his position is that consciously judging that  $p$  gives you justification for believing that you believe that  $p$ .

Silins takes conscious judgements to be ‘those which modify what it is like for you at the time you make them’, so he clearly thinks that forming a judgement has a phenomenology (2012, p. 298). He elaborates:

To make a judgment of the kind I am interested in, you might sincerely assert to someone that  $p$ . But you can consciously judge that  $p$  without performing the linguistic act of assertion. In many cases conscious judgment will be the ‘inner analogue’ of assertion, although it may well be that one can judge that  $p$  without in any way vocalizing or imagining a sentence with the content that  $p$ . (2012, p. 298).

Let us focus for a moment on the words, ‘[i]n many cases conscious judgment will be the “inner analogue” of assertion’. I take it that Silins means the ‘inner analogue’ of *sincere* assertion, given that he explicitly speaks of sincere assertion immediately prior in the case of external speech. I also take it that, by “inner analogue” of assertion’, he means (sincere) assertions made in inner speech, given that, directly afterwards, he contrasts judgements made in this way with judgements made ‘without in any way vocalizing or imagining a sentence’.

Silins motivates his theory by seeking to demonstrate that it can explain why an example of a ‘Moore-paradoxical’ sentence seems irrational. Moore’s Paradox is that there are conjunctions of a particular kind which seem to be irrational to assert or believe even though they are not strictly contradictory. Silins offers this example:

(MP):  $p$  and I do not believe that  $p$ .<sup>91</sup>

Moore’s Paradox is relevant to self-knowledge for the following reason. As just noted, it is not the logical form of a sentence such as MP which makes it seem irrational; the sentence is not actually contradictory. Nor is it anything to do with the content of the sentence; one can substitute any proposition at all for  $p$ . Rather, it seems that the reason that it strikes us as irrational to assert or believe a sentence such as MP is that the sentence expresses a self-ascription of belief that one is plainly not entitled to make.<sup>92</sup> Accordingly, it is a significant advantage for an account of self-knowledge if it can explain what the problem is.

---

<sup>91</sup> Frank Jackson points out to me that it is actually a little strange to think that it is irrational to assert a sentence such as this. Surely this is context-sensitive. If you are offered some reward for asserting MP, then it would be quite rational to assert it.

This raises a challenging question. If it is not necessarily irrational to assert a sentence such as MP, but one thinks that there will still be *something* wrong with asserting MP, no matter the circumstances in which the assertion takes place, then what exactly is it? I will not take up this issue here; I will just proceed on the standard assumption that there is something irrational about asserting (or believing) a sentence such as MP.

<sup>92</sup> Thanks to Luke Roelofs for this. Silins thanks Alan Hájek for alerting him to the relevance of Moore’s Paradox to transparency accounts of self-knowledge.

Silins asks us to focus on *judgements* of the form, MP, and he observes that, '[o]ther things being equal, it is irrational [to] make such judgments' (2012, p. 302). He then seeks to explain the irrationality of these judgements as follows. Suppose you judge that MP is true, i.e. you make the conjunctive judgement:  $p$  and I do not believe that  $p$ . In doing so, you have judged that  $p$  (the first conjunct) and you have judged that you do not believe that  $p$  (the second conjunct). For Silins, though, if you judge that  $p$ , then you have justification to believe that you believe that  $p$ . Accordingly, 'when [you] [judge] the MP conjunction, [you] [judge] the second conjunct, while having justification to believe the negation of the second conjunct' (2012, p. 303).

Silins's account is attractive, but I think there is a problem for it. Silins is explicitly concerned with conscious judgements of the form MP: i.e.  $p$  and I do not believe that  $p$ . And recall that he takes conscious judgements to be judgements 'which modify what it is like for you at the time you make them'. I suggest, though, that there is no way in which judging ' $p$  and I do not believe that  $p$ ' modifies what it is like for you at the time of so judging – because it is simply not possible to judge ' $p$  and I do not believe that  $p$ '. This is so even if (as Silins believes) judging can take the form of producing sincere assertions in inner speech: it is just not possible to sincerely assert in inner speech, ' $p$  and I do not believe that  $p$ '. Of course, you might believe that  $p$  and also believe that you do not believe that  $p$  without being in any way aware of this (though there is, of course, something irrational about this). You might, for example, believe that a particular acquaintance is an unpleasant person. Your behaviour in generally avoiding them would be evidence of this belief. But you might not believe that you believe this person is unpleasant. Perhaps you had formerly been fond of them and do not realise that you have changed your mind.<sup>93</sup> But you cannot consciously judge that this person is unpleasant and that you do not believe that they are unpleasant. (Nor, of course, could you continue to believe that someone is unpleasant, and that you believe that you do not believe that they are unpleasant, after becoming aware that you had been holding this pair of beliefs.) Generally, you cannot consciously judge ' $p$  and I do not believe that  $p$ ' any more than you can consciously judge ' $p$  and not  $p$ '. Moore's Paradox simply does not arise within the context of conscious judgements, because the relevant judgements are just not possible. (It would be quite a paradox if one *could* consciously judge ' $p$  and I do not believe that  $p$ ', notwithstanding the irrationality of the conjunction.)

It is actually odd that Silins sets up Moore's Paradox in terms of conscious judgements. Canonically, the paradox is set up in terms of assertions or beliefs rather than judgements. The original example from Moore involves an assertion. He writes that 'to say such a thing as "I

---

<sup>93</sup> Thanks to Frank Jackson for this point and suggesting an example from which this one derives.



went to the pictures last Tuesday, but I don't believe that I did" is a perfectly absurd thing to say' (Moore 1942, p. 543). And it makes sense that the paradox should be set up in terms of assertions or beliefs. It is not possible to judge ' $p$  and I do not believe that  $p$ ' but it is possible to assert it (though not to assert both conjuncts sincerely), and it is possible to believe both conjuncts (until one becomes aware that one has been doing so). And, of course, it does seem somehow irrational to assert ' $p$  and I do not believe that  $p$ ' or to believe both conjuncts. Why does Supplemented ISI allow us to say about this?

Suppose you believe  $p$ . On Supplemented ISI, if you produce the inner speech assertion,  $p$ , there is something it will feel like for you to do so, and this feeling will justify you in self-ascribing the proposition expressed in the assertion. Now suppose you produce the inner speech assertion, ' $p$  and I do not believe that  $p$ '. We can see why this will seem irrational. Sincerely asserting the first conjunct will justify you in self-ascribing a belief which is the negation of the proposition expressed by the second conjunct. This line of thought clearly follows Silins, but in the context of inner speech assertions rather than judgements.

What about believing both conjuncts in the sentence, ' $p$  and I do not believe that  $p$ ', though without yet realising that you do? What is irrational about this? I suggest the following. Plausibly, rationality requires that sentences expressing all of your beliefs can be conjoined and that the resulting conjunction can be sincerely asserted – in inner speech or external speech. But this will not be possible if your conjunction includes both ' $p$ ' and 'I do not believe that  $p$ '. For as soon as you were to assert  $p$  and have the feeling of having done so sincerely, you would be justified in self-ascribing a belief which is the negation of another of the conjuncts, i.e. 'I do not believe that  $p$ '. This really just combines the argument of the last paragraph with a plausible general principle about rationality.

So Supplemented ISI succeeds where Silins's account fails. Supplemented ISI, rather than Silins's account, can explain the irrationality of sentences such as MP – which, as noted above, is an advantage for a theory of self-knowledge. Clearly, though, Silins's argument lit the way.

Before moving on, it is worth recalling why all of this is relevant to the Vulnerability (IS) Problem. The point is just that Supplemented ISI provides a good account of the apparent irrationality of a Moore-paradoxical sentence, which is a significant advantage for a theory of self-knowledge. It provides one reason to believe that the theory is right in attributing a justificatory role to the phenomenology of producing a sincere assertion in inner speech in self-ascriptions of belief. Accordingly, it provides some support for the central principle which is also appealed to in the answer to the Vulnerability (IS) Problem which Supplemented ISI provides.

### 6.3. The Application of Supplemented ISI

This chapter has focussed on one kind of mental state, belief, and one kind of speech act, assertion. But we self-ascribe many different mental states on the basis of many different kinds of speech acts performed in inner speech. At this point, I just want to suggest that the prospects that Supplemented ISI can be extended to deal with other kinds of mental states and other kinds of speech acts are good.

How it feels to express a wish varies depending whether or not it expresses a genuine desire. How it feels to make an exclamation varies depending whether or not one is genuinely surprised. And this applies in inner speech as much as in external speech. Plausibly, then, there is something it feels like to perform each of these speech acts authentically, and this feeling justifies one in self-attributing the relevant mental state.

## 7. Dialogue Revisited

I now come to a point which has been foreshadowed a couple of times. In Chapter Two, I argued that inner speech is not dialogic. I sought to show that the framework which makes interpersonal dialogue possible – whatever exactly that framework might be – is simply not applicable to inner speech. The example I relied on a number of times related to deception. It is possible (at least in principle) to deceive another person in interpersonal speech, but it is not possible to deceive yourself in inner speech, because there is no distinct interlocutor relying on any expectation that you will speak truthfully.

We saw that a possible objection to this line of reasoning is that the idea that there is no one in inner speech who is liable to be misled assumes that we have more direct access to our mental states than may actually be the case. It assumes, for example, that a subject who believes some proposition,  $p$ , and who produces an utterance in inner speech expressing  $\sim p$ , can know what she believes independently of the inner speech that she produces. This is why it is not possible for her to deceive herself. But some thinkers, such as Carruthers and Cassam, reject this assumption. They argue that we do not have independent access to our mental states; we acquire this by interpreting data such as our inner speech.

I noted that one could respond to the objection simply by pointing out that the assumption is derived from an assumption which features in theories of interpersonal dialogue. For how could one know, for example, that one is speaking truthfully aloud unless one has independent access to one's beliefs? In this chapter, though, I have been developing a view similar to

Carruthers' and Cassam's, on which we do acquire self-knowledge in large part by interpreting our own inner speech. So where does this leave us?

The key point to notice is that, on the terms of Supplemented ISI also, it is not possible to deceive oneself by producing an utterance in inner speech which you know to be false. For what will happen if you produce such an assertion? You will experience the phenomenology characteristic of producing an assertion which you do not believe. *This* will prevent you from deceiving yourself.

Accordingly, the objection raised to the line of argument in Chapter Two can be dismissed whether one thinks that we have direct access to our mental states or whether one thinks that we acquire self-knowledge by, *inter alia*, interpreting our inner speech. If one thinks that we have direct access to our mental states, then one should think that we cannot deceive ourselves simply by producing inner speech assertions which we know to be false. Our direct access to our mental states will allow us to see immediately that we do not believe the assertion. If one thinks that we acquire self-knowledge by, *inter alia*, interpreting our inner speech, and also accepts the supplementation to this view suggested in this chapter, then one should still think that we cannot deceive ourselves simply by producing inner speech assertions which we do not believe. We will experience the wrong phenomenology when we do so.

## **8. Some Evaluation**

My purpose in this chapter has been to propose an amended version of the view that inner speech can be a source of self-knowledge, and to show how it can offer an answer to a problem which arises for the original view. As foreshadowed in the Introduction to the chapter, I have not sought to provide a compelling case for the amended view, but I did offer one argument to establish its plausibility. Still, in closing, it will be worth offering some initial evaluation of the amended view. What follows notes both some strengths and some weaknesses.

### **8.1. Some Strengths**

#### **8.1.1. Reconciling Intuitions**

This chapter began by noting some apparently inconsistent intuitions about self-knowledge. One was that inner speech is an important source of self-knowledge. Another was that our access to our own mental states is different from our access to the mental states of others. Supplemented ISI significantly reduces the tension between these intuitions. On the one hand, it retains an extremely important role for inner speech in self-knowledge. On the other, it holds

that there is a major difference between the access that inner speech provides to our own mental states and the access that external, interpersonal speech provides to the mental states of others. In the case of access to our own mental states, the phenomenology associated with producing a speech act for which all relevant conditions are fulfilled plays a key justificatory role, whereas this is obviously not the case with respect to access to the mental states of others.

### **8.1.2. Uncertainty**

In the discussion so far, I have been assuming, for ease of explication, that we either believe or disbelieve propositions. But things are, of course, not so simple. We are very confident in the truth of some propositions; very uncertain about others. A slightly more sophisticated version of Supplemented ISI could deal with this issue too.

Making an assertion in external speech which one strongly believes feels different from making an assertion in external speech about which one is uncertain. There is a feeling akin to hesitation in the latter case – a feeling of lingering doubt. And so it is with assertions made in inner speech; the phenomenology one experiences when making an assertion will vary according to how confident one is in the truth of that assertion. This phenomenology, in turn, allows one to see how confident they are in the truth of the assertion they make.

On the face of it, this is a very straightforward development of Supplemented ISI. Whether there are complications for it, and whether any such complications could be addressed, will have to be left for further investigation.

### **8.1.3. Actual Speech and Imagined Speech Again**

In Chapter One, I argued that inner speech is a type of actual speech, just as much as external speech is a type of actual speech. Producing inner speech is not the same as imagining speaking. Supplemented ISI comports nicely with this view and perhaps even provides further evidence for it.

If Supplemented ISI is right, then the experience of (say) producing an assertion in inner speech for which all of the relevant conditions are satisfied has a certain qualitative character; the experience of producing an assertion in inner speech for which not all of the relevant conditions are satisfied has a different qualitative character. The same is true when one produces an utterance in actual, external speech, but it is not so when one imagines producing an assertion. If one imagines producing an assertion, then one does not experience the phenomenology of producing an assertion which one either believes or disbelieves, any more

than one does if one produces an assertion in external speech in the context of a theatrical production. This would seem to be another reason to believe that inner speech and imagined speech are different things, and that inner speech is, like external speech, a kind of actual speech.

## **8.2. Some Weaknesses**

### **8.2.1. Assumptions**

In order to be attracted to the supplemented version of Inner Speech Interpretationism and the way it has been presented in this chapter, one needs to share one important background commitment and one important intuition. One needs to believe that inner speech and the interpretation of it plays a role in self-knowledge. And one needs to believe that there is a difference in phenomenology between otherwise identical episodes of producing an assertion depending on whether or not it is sincere.

Both issues can be dealt with quickly. Insofar as Supplemented ISI is intended to offer some defence against a problem which confronts ISI, it is really directed to those who sympathise with ISI, or who would do so if it had some response to that problem. If one never gets on board with ISI, then we need not worry that they will get off.

There is simply not much to be said about the second issue. If someone were to deny that the quality of a subject's experience would vary between otherwise identical episodes of producing an assertion depending on whether or not all of the relevant conditions for the speech act are satisfied, then it is true that nothing could be done to convince them otherwise. It is worth reiterating, though, that it does not matter how one classifies the difference in phenomenology – whether as a difference in cognitive, sensory, or any other kind of phenomenology – so long as one thinks there is some difference in phenomenology.

### **8.2.2. Other Sources of Self-Knowledge**

It has been noted a couple of times in this chapter that a proponent of ISI is likely to have certain other sympathies. One of these is that they are likely to think that we can also learn about our beliefs from our behaviour and from our other conscious, episodic thoughts. This raises the question: can the kind of justificatory account which is offered by Supplemented ISI be extended to explain how we are justified in self-ascribing mental states on the basis of these other sources of sources?

It is very difficult to see how the account could be so extended. Suppose a proponent of ISI also believes that we can sometimes self-ascribe beliefs on the basis of visual imagery. There does not seem to be any room for anything akin to the phenomenology of producing a sincere assertion in inner speech to play a role here. This could be problematic if one wants to be able to tell a consistent story about how we are justified in self-ascribing mental states on the basis of inner speech and on the basis of other sources such as our other conscious, episodic thoughts.

### **8.2.3. Notions of Justification**

The Vulnerability (IS) Problem asks for a justification for a subject's belief that an assertion in inner speech expresses a belief she holds, when this is actually so. Supplemented ISI offers some justification, but not a complete justification. It is open to ask questions such as why it is that we experience a certain phenomenology when, for example, we produce sincere assertions in inner speech, and Supplemented ISI does not even purport to provide an answer. If someone wants an answer to this question, then they will be disappointed with Supplemented ISI.

It is worth noting, of course, that we do not really have a complete justification for any belief – not even for belief in an external world. Still, whether one is satisfied with Supplemented ISI will depend very much on what they want from it.

## Conclusion

The primary objective of this dissertation was to add to the small but growing philosophical literature on inner speech. The three chapters have addressed distinct questions about inner speech, though connections have been noted from time to time. In concluding, I want to show that the positions which have been argued for or (in the case of Chapter Three) explored are consistent. I also want to review some of the respects in which – we have now seen – better understanding inner speech can cast light on other topics, which it was a secondary objective of the dissertation to explore.

The position argued for in Chapter One was that inner speech is a kind of actual speech rather than a kind of imagined speech or a kind of quasi-speech. The position argued for in Chapter Two was that inner speech is not dialogic, in the way that that notion has been explicated. The position that was explored, though not wholly endorsed, in Chapter Three was that the phenomenology of producing a sincere assertion in inner speech can justify self-ascribing belief in the proposition expressed in the assertion. Are these positions compatible?

First, it is certainly the case that inner speech can be a kind of actual speech and not be dialogic. There is a temptation to say that this is obviously true. After all, there can certainly be external speech which is not dialogic. This is the case whenever someone speaks externally and does not have an interlocutor. But, of course, things are a little more complicated than this. Here we should revisit a point which was made at the beginning of Chapter Two.

At first blush, it is tempting to think that inner speech could only be dialogic if it could involve imagining an interlocutor speaking. If this were so, then the finding that inner speech is a kind of actual speech would imply that it is not dialogic. But there is a way to understand the theory that inner speech is dialogic which does not depend on the possibility of imagining interlocutors. On this alternative understanding of the theory – which I suggested should be preferred – dialogue can take place in inner speech when one adopts various perspectives oneself and produces inner speech utterances from them. One might adopt a particular perspective by, for example, imagining having a particular perceptual viewpoint or entertaining a certain set of propositional attitudes.

However, nothing in the argument against the theory that inner speech is dialogic depended on a premise that inner speech is not a kind of actual speech. In a word, the argument was simply that the framework that makes interpersonal dialogue possible – whatever exactly that framework might be – is not transferable to inner speech. And dialogue between individuals, of course, takes place in actual (external) speech.

There is also no tension between the view that inner speech is a kind of actual speech and the view that the phenomenology of sincerely asserting something in inner speech can justify the self-ascription of a belief. Actually, as we saw towards the end of Chapter Three, a key part of the account that I called Supplemented ISI provides support for the view that inner speech is a kind of actual speech. To recap: a central claim of Supplemented ISI is that there is a phenomenological difference between producing an assertion in inner speech sincerely or insincerely – just as there is a phenomenological difference between producing an assertion in external speech sincerely or insincerely. However, there is no such difference between imagining producing a sincere assertion and imagining producing an insincere assertion. This seems like another reason to believe that inner speech is a kind of actual speech (like external speech) rather than a kind of imagined speech.

A second connection between Chapter One and Chapter Three should also be noted. The finding that inner speech is a kind of actual speech is relevant to the treatment of Moore's Paradox in Chapter Three. It is only because inner speech is a kind of actual speech that there seems to be anything irrational about asserting a Moore-paradoxical sentence, such as 'MP' from Chapter Three, in inner speech. Nothing about imagining asserting a sentence such as MP would seem irrational. This recalls the point which was made in setting out the Actual Speech View in Chapter One, that inner speech assertions are truth-evaluable. An individual who produces an assertion in either external speech or inner speech is right or wrong – they speak truly or falsely – but this is not the case if one imagines producing an assertion.

The relationship between Chapters Two and Three is interesting. In Chapter Two, I noted that the evidence for one part of the theory that inner speech is dialogic is strong. This part of the theory is the claim that we represent different perspectives in inner speech. On the interpretation of the theory that I prefer, representing different perspectives involves adopting those perspectives and producing inner speech utterances from them. Adopting a perspective may involve entertaining a set of propositional attitudes and producing inner speech as if one actually holds those propositional attitudes. According to Supplemented ISI, the view developed in Chapter Three, there is something it is like to produce an assertion in inner speech which one believes. So, if an individual entertains a set of propositional attitudes that they do not actually hold, and produces an assertion expressing a proposition within that set, will they experience the phenomenology of producing a sincere assertion?

The answer to that question is straightforward. If the individual simply does not believe the proposition, they will not experience the relevant phenomenology. This is just a case of asserting something you do not believe. But there are some nearby cases in which something more interesting happens. These are cases in which we reason hypothetically – we entertain a



set of premises we do not believe and try to trace out their implications – and realise that we in fact do believe something that follows from, or is suggested by, the premises we have been entertaining. That is, we realise we believe something implied by the premises, even though we do not believe the premises themselves. The realisation can occur upon producing an assertion in inner speech and experiencing the phenomenology of producing a sincere assertion. It can also be quite a surprising experience, especially if the propositions one was entertaining are propositions on which one is not merely neutral, but which one actually rejects.

So, the position that we represent different perspectives in inner speech, and the position that there is something it is like to produce a sincere assertion in inner speech, are not merely consistent. Taken together, they have some additional explanatory power.

In Chapter Two, as well as acknowledging that we can adopt different perspectives in inner speech, I argued that inner speech cannot be self-regulating, borrowing some important material from the philosophy of language to develop this notion. In Chapter Three, as well as suggesting that producing a sincere assertion has a certain phenomenology, I argued that this phenomenology can justify self-ascribing the proposition expressed by the assertion. There does not appear to be any inconsistency between these two positions.

A secondary aim in this dissertation was to show how investigating the phenomenon of inner speech can also shed light on our understanding of other issues. A number of implications of the finding that inner speech is a kind of actual speech were noted at the end of Chapter One. Perhaps the most interesting related to the point that it seems implausible that producing an utterance in inner speech could be morally wrong or blameworthy, no matter how rude, offensive or judgemental that utterance is. Many moral theories hold that it can be wrong to actually say something. Accordingly, if inner speech is a kind of actual speech, then it seems that proponents of these theories owe an explanation as to why inner speech utterances are not apt for moral judgement. At the end of Chapter Two, we saw that the finding that inner speech is not dialogic creates a problem for those who would appeal to some kind of Gricean apparatus in explaining the natural language conditional. We also saw that the reasoning that was employed in arguing against the view that inner speech is dialogic could not successfully be employed in arguing against the view that inner speech solves what Dennett calls ‘accessibility problems’, i.e. facilitates a kind of communication between otherwise separate parts of the brain. The whole discussion of Chapter Three is located within debates about self-knowledge. An appendix will be concerned with a theory that gives a central role to inner speech in seeking to explain auditory verbal hallucinations, or experiences of hearing voices. The discussion there will draw on material from Chapters One and Two.

\* \* \*

At the beginning of this dissertation, I noted that philosophy had neglected inner speech, at least until very recently. I noted there that the discipline had paid a great deal of attention to the nature of the mind and to the nature of speech and language, but not to this phenomenon which lies at their intersection. In the course of this dissertation, I hope that I have shown that the oversight is worth attending to. The phenomenon is of independent interest, and the range of issues which are illuminated if we understand it better is surprisingly large.

## Appendix: Inner Speech and Auditory Verbal Hallucinations<sup>94</sup>

### 1. Introduction

An auditory verbal hallucination ('AVH') is the experience of seeming to hear a voice saying something, despite the absence of any such voice in the external world. AVHs are a key characteristic of schizophrenia but their occurrence is not limited to this condition. They are associated with a range of other psychiatric conditions (such as mood disorders and borderline personality disorder) and neurological conditions (such as epilepsy) (Larøi et al. 2012) and are also known to occur in the general population (Romme et al. 1992; Johns et al. 2002; Sommer et al. 2010). However, it is not known what causes AVHs. A theory which has had significant influence is that AVHs are in fact utterances in a subject's own inner speech which the subject fails to recognise as their own and consequently experiences as coming from without. To put this in the standard terminology, the theory holds that there is a failure of *self-monitoring* (the subject does not recognise the inner speech utterance as their own) leading the subject to *misattribute* the cause of the experience to an external source (e.g. Frith 1992). I will refer to this theory as the 'Inner Speech Theory'.

The purpose of this appendix is to argue for a slightly different account of AVHs, albeit one that still has self-monitoring at its heart. Wayne Wu and Raymond Cho (Wu 2012; Cho & Wu 2013) have suggested that the most plausible version of a self-monitoring account of AVHs would claim that AVHs involve not a failure to monitor the production of one's inner speech, but a failure to monitor the production of auditory imaginings of the voices of others.<sup>95</sup> I will now seek to develop this theory, which I will call the 'Imagined Speech Theory', drawing at a couple of points from Chapters One and Two of this dissertation.<sup>96</sup> I will also discuss some reasons that the Imagined Speech Theory should be preferred over the Inner Speech Theory.

---

<sup>94</sup> The majority of this appendix has been published in the *Review of Philosophy and Psychology* (Gregory 2016).

<sup>95</sup> Although Wu and Cho actually prefer a quite different theory of AVHs. Existing self-monitoring accounts urge that AVHs occur when the intentional production of an inner speech utterance is not properly monitored. On these accounts, an intentionally generated episode is experienced as unintended and an alternative explanation for it is postulated. By contrast, Wu and Cho claim that AVHs are simply experiences of unintended – as they put it, 'automatic' (Wu 2012) or 'spontaneous' (Cho & Wu 2013) – voice-like episodes in the mind. My concern in this appendix is specifically with self-monitoring accounts, so I will not engage with Wu and Cho's own preferred theory.

<sup>96</sup> In Chapter One, I called the theory that inner speech is a kind of imagined speech the 'Imagined Speech View'. Now I am calling the theory that AVHs occur when there is a failure to monitor imagined speech the 'Imagined Speech Theory'. Obviously enough, the similarity in the names results from the central role that imagined speech plays in each theory. It is a little unfortunate, but the two theories should not be confused.

The second section of this appendix will provide an overview of the Inner Speech Theory. The third section will describe one particular feature of AVHs which is highlighted by Wu and Cho: that AVHs very often take place in voices other than the subject's. The fourth section will reiterate a point Wu and Cho have made as to why (what I am calling) the Imagined Speech Theory is better placed to account for this feature of AVHs than the Inner Speech Theory, and it will add to the defence of this line of thought against an objection it has met. The fifth section will itemise some further advantages of the Imagined Speech Theory over the Inner Speech Theory.

One point should be made before commencing. The Inner Speech Theory has largely been developed by theorists concerned with AVHs in schizophrenia. Accordingly, a lot of the research which is reviewed in this Appendix relates to AVHs in schizophrenia. As already noted, though, AVHs do not only occur in schizophrenia. With one exception, the points made in comparing the merits of the Inner Speech Theory and the Imagined Speech Theory are general ones, which do not depend on anything particular to schizophrenia. It has of course occurred to theorists previously that various symptoms in schizophrenia may involve a failure to discriminate what is real from what is imagined. It will be noted in the fifth section that the Imagined Speech Theory coheres neatly with this broader view about schizophrenia.

## **2. The Inner Speech Theory**

The Inner Speech Theory has been advanced or carefully discussed (in whole or part) by a number of authors. The following summary follows quite closely Christopher D. Frith's influential (1992) iteration of it, but also incorporates the contributions of several other authors.

### **2.1. The Basic Version**

The Inner Speech Theory first highlights evidence that there is a connection between AVHs and covert speech generation. Frith refers to evidence including:

1. reports of patients who would simultaneously hear voices and produce whispers, the content of which seemed to correspond to the content of the voices (Gould 1949; Green & Preston 1981);
2. a report of a patient whose 'oral muscle activity' increased immediately prior to episodes of voice-hearing (McGuigan 1966); and
3. experiments finding that auditory hallucinations could be reduced in some patients by 'occupying the speech musculature', either by holding the mouth open wide or

humming (Bick & Kinsbourne 1987; Green & Kinsbourne 1989) (the quotations in these three points are from Frith 1992, pp. 71-72).

Another interesting study (not referred to by Frith) found that increased activity in the speech muscles is more strongly associated with voices which the subject experiences as loud (Inouye & Shimizu 1970).<sup>97</sup>

Even if there is evidence of a connection between inner speech and AVHs, however, this is only part of an explanation. As Frith puts it: '[i]f hallucinations are caused by inner speech, then the problem is not that inner speech is occurring, but that patients must be failing to recognise that this activity is self-initiated. The patients misattribute self-generated actions to an external agent' (1992, p. 73). The suggestion of the Inner Speech Theory is that AVHs occur when a subject's monitoring of their inner speech fails, so that they produce inner speech but do not recognise it as their own.

This suggestion is often illustrated by pointing to an example of self-monitoring in another sensory domain – the visual domain – and showing how failures of this self-monitoring can lead to anomalous experiences (e.g. Frith 1992, Langland-Hassan 2008). When we move our eyes, we do not experience the world as moving around us; rather, we experience the world as remaining still while our gaze moves across it. The explanation is that we monitor the movement of our eyes: when an instruction to move is sent to the eye muscles, something like a copy of this instruction, often called a 'corollary discharge', is sent to a monitoring system in the brain, so that a change in the visual field is expected. Accordingly, the change in the visual field which occurs is attributed to one's own movement, rather than a movement in the environment, which is experienced as stationary. But the system can be interfered with. For example, if an eye is caused to move other than by the muscles which usually control it (e.g. by moving it gently with a finger), then no appropriate message will be sent to the monitoring system. The change in the visual field which results from the movement of the eye is experienced as a movement of the surrounding environment (Frith attributes these observations to Helmholtz 1866).

The Inner Speech Theory then highlights evidence that subjects who experience AVHs may have difficulties with self-monitoring generally. Among other things, Frith cites two separate studies in which subjects were required to use a computer joystick to follow targets on a display. The movements of the targets often induced subjects to move the joystick in an incorrect direction, but subjects can usually correct errors of this kind prior to receiving visual feedback (Frith 1992 cites Megaw 1972, Rabbitt 1966). Frith explains: '[i]t is argued that the

---

<sup>97</sup> See Footnote 6. It is with respect to the sources of evidence just reviewed that the notion that inner speech involves activity in the speech production muscles is relevant.

ability of people to make these very rapid error corrections in the absence of external feedback demonstrates that they are monitoring the response intended (via corollary discharge) and thus do not need to wait for external feedback' (1992, p. 83). Frith cites Malenka et al. (1982), however, as finding that subjects with schizophrenia were less likely than other subjects to correct their errors without visual feedback. He also cites Frith & Done (1989) as confirming this but also finding that it was only patients who have 'passivity experiences' who were unable to adjust their actions without visual feedback (Frith 1992, p. 83). For Frith, this is significant as passivity experiences – i.e. 'symptoms of schizophrenia that explicitly concern the attribution of the patient's own actions to outside agents' (1992, p. 73) – are the very symptoms which are amenable to be explained by a failure of self-monitoring.

So, there is some evidence of a connection between internal, linguistic activity and AVHs, and some evidence that subjects with schizophrenia – and especially those whose symptoms include passivity experiences – have difficulties with self-monitoring. The suggestion, then, is that AVHs occur when a subject produces an inner speech utterance but somehow does not properly monitor the production of this inner speech utterance. The subject then has a sensory experience – seeming to hear something being said – which they do not recognise as self-generated.

## **2.2. Self-Monitoring and Agency**

This basic version of the Inner Speech Theory can be enhanced by incorporating a suggestion from Irwin Feinberg. In an important paper, Feinberg (1978) speculated about how corollary discharge mechanisms might explain the feeling of agency in conscious thinking. Feinberg first noted John Hughlings Jackson's (1958) suggestion that conscious thinking may be a sensory-motor process, presumably having evolved from other, more basic sensory-motor processes. Feinberg then considered what would follow if corollary discharges were generated when we produce conscious thoughts, and if these corollary discharges were also conscious. He suggested that 'the subjective experience of these discharges should correspond to nothing less than the experience of will or intention' (Feinberg 1978, p. 628). His point seems to be that the experience of these discharges would be confirmation for an agent that they were the source of the conscious thoughts occurring in their mind. Feinberg went on to suggest that the malfunctioning of this process could be what accounts for some of the experiences which are

characteristic of schizophrenia – including AVH, in which one has a conscious experience but does not recognise it as self-generated.<sup>98</sup>

John Campbell (1999) suggested that Feinberg’s model could be improved by dropping the notion that the corollary discharges accompanying conscious thoughts are themselves conscious. For if they were, then we might expect ‘a “double awareness” of [the] contents [of our conscious thoughts], once as the contents known by introspection, and once as the contents of the [corollary discharges]’ (Campbell 1999, p. 618). So how might corollary discharges explain the sense that one is the agent of one’s thoughts, if corollary discharges are not conscious? Campbell adopts and extends Georgieff & Jeannerod’s (1998) suggestion that the sense of agency which accompanies physical actions derives from the match, which is detected subconsciously by a ‘comparator’ in the brain, between the corollary discharges which issue when one performs those actions and one’s sensory experience of those actions. For example, if you move your arm before your face, ‘[w]hat explains the feeling that it is you who moved your arm is that at the comparator, [a corollary discharge] was received of the instruction to move which matches the movement you perceive’ (Campbell 1999, p. 612). Campbell proposes that the same applies to conscious thoughts. What accounts for our experience of agency in our conscious thinking is the match, detected subconsciously, between our experience of our conscious thoughts and the corollary discharges which issue when those thoughts are generated – not conscious experience of the corollary discharges themselves.

The basic version of the Inner Speech Theory sketched above suggests that AVHs involve a failure in monitoring one’s own inner speech. What Feinberg’s and Campbell’s work offers is a promising, albeit schematic account as to how self-monitoring might generally account for our experience of agency in our own occurrent thoughts – and therefore as to why that sense of agency might go missing when self-monitoring fails.

### **2.3. Further Developments**

Recent work has developed the Inner Speech Theory in different directions. Seal et al. (2004) propose that a failure of self-monitoring will cause a subject to experience an inner speech utterance as unintended and therefore to feel some uncertainty as to its origin. They suggest that ‘top-down’ factors such as current mood and an ‘externalising response bias’ then cause the subject to attribute the utterance to an external source (2004, p. 60). By contrast, Jones & Fernyhough (2007) suggest that a failure of the self-monitoring of inner speech would prevent

---

<sup>98</sup> Frith does mention Feinberg’s paper briefly (1992, p. 81). See Frith 1992, pp. 80-93, generally for discussion of the monitoring of intentions and for references to further literature.

the ‘emotion of self-authorship’ (p. 393) which ordinarily attends the production of occurrent thoughts from arising, so that an inference to ‘other authorship’ (pp. 394, 395) would be automatic (Jones & Fernyhough cite Wegner (2002) for the term ‘emotion’ of self-authorship). That is, they deny there is any intervening uncertainty before the subject imputes an external cause for their experience.

Referring to Jones & Fernyhough (2007) and also to Proust (2006), Peter Langland-Hassan (2008) questions whether there is any distinct emotion of ‘self-authorship’ which ordinarily accompanies conscious thoughts (p. 391; see also p. 371, citing Strawson 2003) and also proposes a model of self-monitoring which does not require a comparator mechanism. First, Langland-Hassan emphasises the fact that self-monitoring of our own actions causes us to have an ‘attenuated’ experience of them (2008, p. 397). For example, as Langland-Hassan (citing Weiskrantz et al. 1971) notes, most people are unable to tickle themselves: we are not as sensitive to our own touch as we are to that of others. Langland-Hassan then proposes that this attenuation of sensitivity to our own actions might not require any comparison between expected and actual sensory input. Rather, it might be that the corollary discharges which issue when we perform physical actions allow us to anticipate certain sensory input, and therefore to ‘filter’ this input when it occurs (2008, p. 383).

For Langland-Hassan, the sense of agency that we ordinarily have in producing our own actions is just ‘embedded’ (2008, pp. 391, 392) in the way that we experience those actions – i.e. in our attenuated sensory experience of them. It is not a separate, ‘inscrutable emotion’ (2008, p. 392). Langland-Hassan argues that our experience of our own inner speech is very likely attenuated in the same way that our experience of our own external actions is, and suggests that this may be what ‘imbue[s]’ (2008, p. 384) it with a sense of agency. On this proposal, it is when this attenuation fails that one would have an unusual experience of their inner speech, which might then be interpreted as an external voice.

#### **2.4. Putting it Together**

From all of the above, we can piece together the following picture. We monitor the performance of our own physical actions. When we perform physical actions, corollary discharges issue, predicting changes in the sensory feedback which we receive. When there is a match between corollary discharge and sensory feedback, the action is experienced as a consequence of one’s own intentions. This may also be true of the generation of conscious, episodic thoughts, including inner speech utterances. If a subject produces an inner speech utterance but does not properly monitor the production of this utterance, then the subject will



not experience it as originating from him/herself and will postulate an external cause for it. That is, the subject will experience an AVH.<sup>99</sup>

Soon, we will see how this picture faces a significant problem. First, though, it is necessary to say a little bit more about the phenomenology of AVHs.

### 3. Hearing Voices

There is great diversity in experiences of voice-hearing. As Frank Larøi (2006) observes, one study (Stephane et al. 2003) of AVHs in a sample of 100 psychiatric patients explored subjects' experiences with respect to some 20 different variables, which gives a sense of the heterogeneity of the phenomenon. There is one particular attribute which is common in AVHs, however, and which should raise some doubt about the Inner Speech Theory. As both Wu (2012) and Cho & Wu (2013) observe, AVHs very often do not take place in the voice of the subject who is experiencing them. They take place in the voices of others. This is apparent from some clear themes in the empirical literature.

First, subjects can often indicate the identity of voices that they hear. In Nayani & David's (1996) study of 100 subjects with psychosis, 61 could nominate the identity of at least some of their voices. The identifications were 'delusional' in some instances (such as the voice of God or the Devil) and 'real/likely-to-be-real' in others (examples given are 'relative, neighbour, doctor') (1996, p. 181). In Leudar et al.'s (1997) study of a sample of voice-hearers both with and without diagnoses of schizophrenia, 10 out of 25 subjects reported hearing voices which were 'aligned' with (i.e. seemed similar to) family members; 4 out of 25, voices aligned with acquaintances; 6 out of 25, voices aligned with public figures; and 5 out of 25, voices aligned with supernatural figures (pp. 890-891). And in Sommer et al.'s (2010) study, 58% of 103 non-clinical voice-hearers 'attributed the voices to an external source, mostly benevolent spirits' (p. 635).

The evidence in the last paragraph is particularly direct. That AVHs regularly take place in voices other than the subject's own voice, however, can also be readily inferred from other aspects of the accounts voice-hearers provide. For example, subjects who experience AVHs

---

<sup>99</sup> It is often suggested that there might be no single, unified explanation for AVHs. Frith, for example, offered his theory as an explanation for only 'certain' AVHs in schizophrenia (Frith 1992, pp. 84, 115). He thought it was most promising for cases in which subjects seem to hear their 'own thoughts spoken aloud' (1992, p. 84) (though he elsewhere labels such experiences '[t]hought broadcast' (1992, p. 66) and we might doubt that they are really AVHs as usually understood). As we will see, though, one reason to prefer the Imagined Speech Theory over the Inner Speech Theory is the very fact that it is applicable to a larger range of cases. It would not be an answer to such an argument to assert that the Inner Speech Theory is only intended to have limited application.

report frequently that they hear multiple voices. Nayani & David (1996) reported that their subjects heard a mean of 3.2 voices. In McCarthy-Jones et al.'s (2014) study of almost 200 psychiatric patients, subjects heard an average of 4.3 voices (even after those who heard an 'uncountable' number of voices were excluded) (pp. 228, 232). In Leudar et al.'s (1997) study, 89% of subjects experienced particular voices repeatedly, and the average number of these 'recurrent voices' was 2.5 (p. 890). If subjects generally hear multiple voices, then most of those voices cannot be their own.

It is also common for subjects to report hearing a number of voices speaking to one another (Nayani & David 1996; Sommer et al. 2010; McCarthy-Jones et al. 2014; though cf. Leudar et al. 1997, especially p. 892). Even if one of those voices is the subject's own, the voices with whom it is conversing must still be those of others. It is also clear that some subjects' voices have attributes not present in the subjects' own voices. For example, both Nayani & David (1996) and McCarthy-Jones et al. (2014) found that subjects' voices were more often male voices – even for female subjects.

Some subjects do experience AVHs in their own voices. For example, six of 25 subjects in Leudar et al.'s (1997) study reported this and Sedman (1966) described some patients who reported hearing their own voices. However, it is clear that very many AVHs take place in the voices of others. (Wu (2012) also notes some of this literature, and makes the point about the apparent gender of voices.)

#### **4. A Problem for the Inner Speech Theory**

Cho & Wu (2013) and Wu (2012) raise a problem for what I am calling the Inner Speech Theory which highlights a mismatch between the typical experience of AVHs as involving voices other than one's own and the experience that we might expect misattributed inner speech to bring about. The particular problem that they raise, however, is also suggestive of an alternative self-monitoring account of AVHs.

##### **4.1. Inner Speech and Imagined Speech Again**

The problem which Cho & Wu (2013) and Wu (2012) raise for what I am calling the Inner Speech Theory is most easily demonstrated with the aid of an example. Take a case in which you have an experience as of someone else speaking. Perhaps you have an experience as of a friend saying, 'Don't forget the match starts at 8', without that friend being present. If this mental episode occurs in a non-hallucinatory context – i.e. if you experience the episode as a

product of your own agency – would we think of it as an instance of inner speech? Cho and Wu are of the view that the answer is no; we would think of it as an instance of imagining someone else speaking (though they put the point in different terminology). This point that imagining someone else speaking is not what we would ordinarily think of as inner speech, of course, also played a part in my discussion in Chapter One regarding whether inner speech is a kind of imagined speech or a kind of actual speech.

Now, still following Cho and Wu’s reasoning, suppose you have a hallucination of a friend saying, ‘Don’t forget the match starts at 8’. Could this result from a failure of monitoring of inner speech, given that the hallucination takes place in the voice of another person? The answer which the Inner Speech Theory provides is that, if one produces an inner speech utterance but (owing to some failure of self-monitoring) does not experience it as their own, then they will experience it as coming from without – i.e. they will misattribute it. But subjects very often do not just experience the utterances as coming from without; they experience them in the voices of other people. Sometimes voices are experienced as issuing from ‘anonymous’ speakers (Leudar et al. 1997, p. 891) but, very often, they are experienced as issuing from particular, identifiable individuals, as in our example. We saw extensive evidence of this in the last section.

So what exactly is going on? Cho and Wu suggest the following explanation as the best possible version of a self-monitoring account of AVHs.<sup>100</sup> AVHs are not instances of inner speech where there has been a failure of self-monitoring but instances of imagined speech where there has been a failure of self-monitoring. For, as I observed in Chapter One, imagined speech certainly can take place in the voices of others. Moreover, if someone imagined some such episode of speech but lost the sense that they were the agent bringing about the act of imagining, then they would presumably experience the voice as coming from without. This is the Imagined Speech Theory.

Before considering an objection to the Imagined Speech Theory and also listing its merits, I want to make one point. The conclusion from Chapter One of this dissertation helps us to better understand the problem which arises for the Inner Speech Theory but not the Imagined Speech Theory. Inner speech, I argued in Chapter One, is a kind of actual speech, just as much as external speech is a kind of actual speech. Supposing this is so, it allows us to see why inner speech only takes place in one’s own voice and not in the voices of others. We only produce external speech in our own voices. If inner speech is just a silent version of external speech,

---

<sup>100</sup> Though they do not endorse this account. See Footnote 96.

then the same will be true of inner speech as well. So, Cho and Wu show *that* there is a problem for the Inner Speech Theory; Chapter One of this dissertation explains *why* it arises.

#### 4.2. An Objection

In response to Cho and Wu, Peter Moseley & Sam Wilkinson (2013) have argued that the Inner Speech Theory can account for the fact that AVHs often take place in the voices of others because inner speech itself can involve the voices of others. Moseley & Wilkinson point to a survey conducted by McCarthy-Jones & Fernyhough (2011) which, Moseley & Wilkinson claim, found ‘that it is common for healthy, non-clinical participants to report hearing other voices as part of their inner speech, as well as to report their inner speech taking on the qualities of a dialogic exchange’ (p. 1).<sup>101</sup> Moseley & Wilkinson suggest that these findings are ‘consistent with Vygotskian explanations of the internalization of external dialogs during psychological development’ (2013, p. 2, citing Fernyhough 2004). This is the idea, explored at length in Chapter Two, that inner speech develops as the child internalises the practice of participating in dialogues with other people – and, possibly, retains some dialogic qualities. Some further, albeit relatively weak evidence, which Moseley & Wilkinson do not rely upon, is available from Hurlburt et al. (2013), who take it that inner speech can take place in the voices of others but that this happens only very rarely. They write: ‘We have seen, over the course of 30 years of investigations, only a handful of such instances’ (2013, p. 1488).

Cho and Wu have themselves replied to Moseley & Wilkinson (Cho & Wu 2014) and, as they note, the detail of McCarthy-Jones & Fernyhough’s survey is important. It elicited responses to such items as the following:

‘I hear the voice of another person in my head. For example, when I have done something foolish I hear my mother’s voice criticising me in my mind’;

‘I experience the voices of other people asking me questions in my head’;

‘I hear other people’s voices nagging me in my head’;

‘I hear other people’s actual voices in my head, saying things that they have never said to me before’; and

‘I hear other people’s actual voices in my head, saying things that they actually once said to me’. (McCarthy-Jones & Fernyhough 2011, p. 1589).

There is a major problem in the way that McCarthy-Jones and Fernyhough interpret the results of their own experiment. As Cho & Wu point out, nothing about a positive response to these

---

<sup>101</sup> See Chapter Two, Section 3.2 for discussion of this study.

items entails that the relevant voices occur in inner speech. On the contrary, it is perfectly consistent with the wording of these items that the voices occur in imagined speech.

But there is more that can be said in reply to Moseley & Wilkinson. In Chapter Two, I argued at length against the theory that inner speech is dialogic, but let us set the conclusion aside. Even if one accepts the Vygotskian picture about the development of inner speech, and also believes that inner speech retains dialogic qualities when it is fully developed, this does not mean that inner speech must involve the voices of others.<sup>102</sup> In the study just considered, McCarthy-Jones & Fernyhough describe dialogic inner speech as involving an ‘ongoing interplay between different internalized perspectives’ (2011, p. 1587), but this obviously does not entail that each perspective is expressed by a different voice. It could simply be that multiple perspectives are expressed by one voice (see also Hurlburt et al. 2013, citing Fernyhough 2004, for this point). This is the interpretation of the idea that one represents different perspectives in inner speech which, I suggested in Chapter Two, should be preferred.<sup>103</sup>

Once these points are noted, there seems to be little reason to think that inner speech really does take place in voices other than one’s own. There is still the handful of instances referred to by Hurlburt et al. These may have carried some weight as part of a cumulative case. On their own, though, they are surely best set aside as anomalous, or possibly as cases of subjects misreporting their own experiences.

## **5. Advantages of the Imagined Speech Theory**

The Imagined Speech Theory has three advantages over the Inner Speech Theory. First, as Wu and Cho argue, the Imagined Speech Theory offers a more complete account of AVHs with fewer resources. On the Inner Speech Theory, it is not clear why subjects should very often experience AVHs in the voices of particular others. Even if a subject does fail to monitor their inner speech, why should they then attribute the utterance which they experience to a family member, friend, public figure or supernatural entity? On the Imagined Speech Theory, however, it is natural that AVHs should often be experienced in the voices of others. If the mental state which is not properly monitored is an episode of imagining some particular

---

<sup>102</sup> Of course, Moseley and Wilkinson do not claim that it does; they just note that the presence of the voices of others in inner speech would be compatible with the Vygotskian picture.

<sup>103</sup> This is all consistent with the view I argue for in Chapter Two, that multiple perspectives are represented in inner speech, but that the episodes of inner speech which Fernyhough and others would consider ‘dialogic’ are not self-regulating in the way that interpersonal dialogues are. The point is just that, even if one subscribes to the theory that inner speech is dialogic, this does not imply that inner speech involves voices other than one’s own.

individual speaking, then we might expect the resulting experience to be as of hearing that individual actually speaking. On the Imagined Speech Theory, then, misattribution is not a separate feature of AVHs which needs to be explained in addition to a failure of self-monitoring; it is a natural consequence of it.

Secondly, the Imagined Speech Theory has more explanatory power than the Inner Speech Theory. Richard Bentall (1990) has suggested that a significant problem ‘for theories linking auditory hallucinations to subvocalization is that they cannot account for hallucinations in the non-auditory modalities’ (p. 88). That is, even if AVHs could be explained as misattributed inner speech, there is no way that the visual, tactile, olfactory etc. hallucinations which are also familiar to many people who experience AVHs could also be explained as misattributed inner speech. Further to Bentall’s point, it also seems implausible that any version of the Inner Speech Theory could account for multi-modal hallucinations (e.g. seeming to see somebody and hear them speak).

By contrast, the Imagined Speech Theory could plausibly be developed to account for both non-auditory hallucinations and multi-modal hallucinations. It may be that hallucinations generally are instances of imagination which are not properly self-monitored. We can certainly have episodes of imagination which are non-auditory and multi-modal. If such episodes were not properly self-monitored, then the corresponding kinds of hallucination are exactly what we might expect to result. (See Cho & Wu (2014) for a related point about auditory *non-verbal* hallucinations.)

Thirdly, the Imagined Speech Theory coheres very neatly with a natural way of thinking about a range of the symptoms of schizophrenia: that they involve a failure to discriminate what is real from what is imagined.<sup>104</sup> A recent proponent of this idea is Gregory Currie (2000), though he notes that the idea is not at all new, singling out the psychiatrist, Harold Searles, as one earlier advocate (Currie 2000, citing Sass 1994, who quotes from Searles). It is easy to see why this way of thinking about symptoms in schizophrenia is attractive. As Currie puts it, ‘much of what exemplifies the strange and disordered thought of people with schizophrenia would not be remarkable if it were treated by the subject as belonging to the flow of her own imaginings’ (2000, p. 174).

The Imagined Speech Theory fits very neatly into this picture: Currie actually mentions auditory hallucinations, including voice-hearing, as an example of the sensory hallucinations which he takes to be explainable on the model he is advancing (2000, p. 180). More generally, though, if one is inclined to think that some symptoms of schizophrenia likely involve a failure

---

<sup>104</sup> This is the one point specific to schizophrenia foreshadowed in the Introduction to this Appendix.

to recognise episodes of imagination as such, then it is easy to see that one already has to hand the resources for at least a general explanation of AVHs. By contrast, the Inner Speech Theory attempts to explain AVHs by recruiting a phenomenon which I endeavoured to show in Chapter One is entirely separate from imagination, i.e. inner speech.

In reflecting on the development of models invoking the comparison of corollary discharges with sensory feedback for explaining various symptoms in schizophrenia and on challenges facing those models, Frith (2012) recently wrote the following:

‘while a general theory of symptoms seems appropriate since some commonality between symptoms is required to justify the persistence of the idea of a unitary diagnosis of schizophrenia, there also needs to be some account of why specific symptoms differ so markedly from one patient to another’ (2012, p. 53).

Importantly, what the Imagined Speech Theory contributes to a unified explanation of symptoms in schizophrenia does not come at the expense of this wide applicability (i.e. the explanatory power noted above). Rather, it is just because the Imagined Speech Theory gives a central place to imagination in explaining AVHs, and because imagination itself permits of such great variety, that the Imagined Speech Theory is able to account for the many different ways in which AVHs are experienced.

One way in which the Imagined Speech Theory is not disadvantaged with respect to the Inner Speech Theory should also be noted. It is not obvious why subjects who experience AVHs should be producing the imagined speech which is not properly monitored, especially where that imagined speech is distressing. However, there is a corresponding problem for the Inner Speech Theory. Even if AVHs are misattributed inner speech utterances, it is not obvious why subjects are producing the inner speech which is not properly monitored, especially where that inner speech is distressing. This is a challenge for both theories.<sup>105</sup>

## 6. Conclusion

I conclude with a word of speculation as to the source of error in the Inner Speech Theory. The question whether AVHs are more likely instances of misattributed inner speech or misattributed imagined speech suggests itself as soon as one distinguishes between inner speech and imagined speech. Probably, those who have taken AVHs to be instances of

---

<sup>105</sup> One possibility available to the Imagined Speech Theory deserves mention. There is some association between the experience of AVHs and a history of childhood trauma in both clinical and non-clinical populations (Sommer et al. 2010; Daalman et al. 2012). It may be worth investigating whether the experience of childhood trauma could contribute to the development of a more active imagination (or to difficulties in controlling one’s imagination). Cf. Cho & Wu (2014) comparing the merits of their preferred theory as against self-monitoring theories regarding this issue.

misattributed inner speech have just failed to draw this distinction. This would appear to be the case in McCarthy-Jones & Fernyhough's (2011) study. As Cho & Wu (2014) rightly point out, subjects' testimony that they have (non-hallucinatory) experiences as of others' voices does not mean that these experiences are cases of inner speech. If one does not distinguish between inner speech and imagined speech, and takes experiences as of others' voices to be a kind of inner speech, then one could well come to the view that AVHs are instances of misattributed inner speech. Accordingly, the error in the Inner Speech Theory most likely results from a failure to analyse the phenomenon of inner speech with sufficient rigour.



## Reference List

- Adler, J 2015, 'Epistemological problems of testimony', in EN Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), viewed 24 May 2017 <<http://plato.stanford.edu/archives/sum2015/entries/testimony-episprob/>>.
- Alderson-Day, B, McCarthy-Jones, S, Bedford, S, Collins, H, Dunne, H, Rooke C & Fernyhough, C 2014, 'Shot through with voices: dissociation mediates the relationship between varieties of inner speech and auditory hallucination proneness', *Consciousness and Cognition*, vol. 27, pp. 288-296.
- Alderson-Day, B, Weis, S, McCarthy-Jones, S, Moseley, P, Smailes D & Fernyhough C 2016, 'The brain's conversation with itself: neural substrates of dialogic inner speech', *Social, Cognitive and Affective Neuroscience*, vol. 11, no. 1, pp. 110-120.
- Atkinson, JR 2006, 'The perceptual characteristics of voice-hallucinations in deaf people: insights into the nature of subvocal thought and sensory feedback loops', *Schizophrenia Bulletin*, vol. 32, no. 4, pp. 701-708.
- Atkinson, JR, Gleeson, K, Cromwell, J & O'Rourke, S 2007, 'Exploring the perceptual characteristics of voice-hallucinations in deaf people', *Cognitive Neuropsychiatry*, vol. 12, no. 4, pp. 339-361.
- Austin, JL 1962, *How to do things with words*, OUP, Oxford.
- Baddeley, A 1992, 'Working memory', *Science*, vol. 255, no. 5044, pp. 556-559.
- Bakhtin, MM 1975/1981, 'Discourse in the novel', C Emerson & M Holquist (trans.), in M Holquist (ed.), *The dialogic imagination: four essays by M. M. Bakhtin*, University of Texas Press, Austin, pp. 259-422.
- Bayne, T & Montague, M (eds.) 2011a, *Cognitive phenomenology*, OUP, Oxford.
- Bayne, T & Montague, M 2011b, 'Cognitive phenomenology: an introduction', in T Bayne & M Montague (eds.), *Cognitive phenomenology*, OUP, Oxford, pp. 1-34.
- Barresi, J & Moore, C 1996, 'Intentional relations and social understanding', *Behavioral and Brain Sciences*, vol. 19, pp. 107-154.
- Bentall, RP 1990, 'The illusion of reality: A review and integration of psychological research on hallucinations', *Psychological Bulletin*, vol. 107, no 1, pp. 82-95.
- Bick, PA & Kinsbourne, M 1987, 'Auditory hallucinations and subvocal speech in schizophrenic patients', *American Journal of Psychiatry*, vol. 144, pp. 222-225.
- Blachowicz, J 1997, 'The dialogue of the soul with itself', *Journal of Consciousness Studies*, vol. 4, no. 5-6, pp. 485-508.
- Brandom, RB 1994, *Making it explicit: reasoning, representing, and discursive commitment*, Harvard University Press, Cambridge, MA.
- Bransford, J & Johnson, M 1972, 'Contextual prerequisites for understanding: some investigations of comprehension and recall', *Journal of Verbal Learning and Verbal Behavior*, vol. 11, pp. 717-726.
- Burge, T 1993, Content preservation, *Philosophical Review*, vol. 102, no. 4, pp. 457-488.
- Campbell, J 1999, 'Schizophrenia, the space of reasons, and thinking as a motor process', *The Monist*, vol. 82, pp. 609-625.

- Carruthers, P 2009 'How we know our minds: the relationship between mindreading and metacognition', *Behavioral and Brain Sciences*, vol. 32, pp. 1-18.
- Carruthers, P 2011, *The opacity of mind: an integrative theory of self-knowledge*, OUP, Oxford.
- Cassam, Q 2014, *Self-knowledge for humans*, OUP, Oxford.
- Cho, R & Wu, W 2013, 'Mechanisms of auditory verbal hallucination in schizophrenia', *Frontiers in Psychiatry*, vol. 4, no. 155, pp. 1-8.
- Cho, R & Wu, W 2014, 'Is inner speech the basis of auditory verbal hallucinations in schizophrenia?' *Frontiers in Psychiatry*, vol. 5, no. 75, pp. 1-3.
- Chuang-tzū 1981, *The seven inner chapters and other writings from the book Chuang-tzū*, AC Graham (trans), Allen & Unwin, London.
- Chudnoff, E. (2013) Intellectual gestalts, in T. Horgan & U. Kriegel (eds.), *Phenomenal intentionality: new essays*, Oxford: Oxford University Press.
- Clowes, R 2007, 'A self-regulation model of inner speech and its role in the organisation of human conscious experience', *Journal of Consciousness Studies*, vol. 14, no. 7, pp. 59-71.
- Currie, G 1995, 'Imagination and simulation: aesthetics meets cognitive science', in M Davies & T Stone (eds), *Mental simulation: evaluations and applications*, Blackwell, Oxford, pp. 151-169.
- Currie, G 2000, 'Imagination, delusion and hallucinations', *Mind & Language*, vol. 15, no. 1, pp. 168-183.
- Daalman, K, Diederens, KMJ, Derks, EM, van Lutterveld, R, Kahn, RS & Sommer, EC 2012, 'Childhood trauma and auditory verbal hallucinations', *Psychological Medicine*, vol. 42, pp. 2475-2484.
- Dennett, DC 1969, *Content and consciousness*, Routledge & Kegan Paul, London.
- Dennett, DC 1991, *Consciousness explained*, Back Bay Books, New York.
- Dennett, DC 1992, 'The self as the center of narrative gravity', in FS Kessel, PM Cole & DL Johnson (eds.), *Self and consciousness: multiple perspectives*, Erlbaum, Hillsdale, NJ, pp. 103-115.
- Dickens, C 1857/1967, *Little Dorrit*, Penguin Books, London.
- Evans, G 1982, *The varieties of reference*, OUP, Oxford.
- Faulkner, P 2007, 'On telling and trusting', *Mind*, vol. 116, pp. 875-902.
- Feigenbaum, P 1992, 'Development of the syntactic and discourse structures of private speech', in RM Diaz & LE Berk (eds.), *Private speech: from social interaction to self-regulation*, Erlbaum, Hillsdale, NJ, pp. 181-198.
- Feinberg, I 1978, 'Efference copy and corollary discharge: implications for thinking and its disorders', *Schizophrenia Bulletin*, vol. 4, pp. 636-640.
- Fernyhough, C 1994, *Social and private speech as determinants of early cognitive functioning*, unpublished doctoral dissertation, University of Cambridge.
- Fernyhough, C 1996, 'The dialogic mind: a dialogic approach to the higher mental functions', *New Ideas in Psychology*, vol. 14, no. 1, pp. 47-62.

- Fernyhough, C 2004, 'Alien voices and inner dialogue: towards a developmental account of auditory verbal hallucinations', *New Ideas in Psychology*, vol. 22, pp. 49-68.
- Fernyhough, C 2008, 'Getting Vygotskian about theory of mind: mediation, dialogue, and the development of social understanding', *Developmental Review*, vol. 28, pp. 225-262.
- Fernyhough, C 2009, 'Dialogic thinking', in A Winsler, C Fernyhough & I Montero (eds.), *Private speech, executive functioning, and the development of verbal self-regulation*, CUP, Cambridge, pp. 42-52.
- Frith, CD & Done, DJ 1989, 'Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action', *Psychological Medicine*, vol. 19, pp. 359-363.
- Frith, CD 1992, *The cognitive neuropsychology of schizophrenia*, Erlbaum, Hove, UK.
- Frith, CD 2012, 'Explaining delusions of control: the comparator model 20 years on', *Consciousness and Cognition*, vol. 21, pp. 52-54.
- Gelfert, A 2015, 'Inner speech, natural language, and the modularity of mind', *Kairos*, vol. 14, pp. 7-29.
- Georgieff, N & Jeannerod, M 1998, 'Beyond consciousness of external reality: a 'Who' system for consciousness of action and self-consciousness', *Consciousness and Cognition*, vol. 7, pp. 465-477.
- Gerrans, P 2015, 'The feeling of thinking: sense of agency in delusions of thought insertion', *Psychology of Consciousness: Theory, Research, and Practice*, vol. 2, no. 3, pp. 291-300.
- Gibbard, A 1990, *Wise choices, apt feelings: a theory of normative judgment*, Harvard University Press, Cambridge, MA.
- Goudena, PP 1992, 'The problem of abbreviation and internalization of private speech', in RM Diaz & LE Berk (eds.), *Private speech: from social interaction to self-regulation*, Erlbaum, Hillsdale, NJ, pp. 215-224.
- Gould, LN 1949, 'Auditory hallucinations and subvocal speech', *Journal of Nervous and Mental Disease*, vol. 109, pp. 418-427.
- Green, MF & Kinsbourne, M 1989, 'Auditory hallucinations in schizophrenia: does humming help?' *Biological Psychiatry*, vol. 25, pp. 633-635.
- Green, P & Preston, M 1981, 'Reinforcement of vocal correlates of auditory hallucinations by auditory feedback: a case study', *British Journal of Psychiatry*, vol. 139, pp. 204-208.
- Gregory, D 2016, 'Inner speech, imagined speech, and auditory verbal hallucinations', *Review of Philosophy and Psychology*, vol. 7, pp. 653-673.
- Gregory, D 2017, 'Is inner speech dialogic?', *Journal of Consciousness Studies*, vol. 24, no. 1-2, pp. 111-137.
- Grice, P 1975/1989, 'Logic and conversation', in P Grice 1989, *Studies in the way of words*, Harvard University Press, Cambridge, MA. Reprinted from P Cole & JL Morgan (eds.) 1975, *Syntax and semantics*, vol. 3, Academic Press, New York, pp. 41-58.
- Grice, P 1989, 'Indicative conditionals', in P Grice, *Studies in the way of words*, Harvard University Press, Cambridge, MA.
- Heavey, CL & Hurlburt, RT 2008, 'The phenomena of inner experience', *Consciousness and Cognition*, 17, pp. 798-810.
- Helmholtz, H 1866, *Handbuch der physiologischen Optik*, Voss, Leipzig.

- Hobson, RP 1995, 'Apprehending attitudes and actions: separable abilities in early development', *Development and psychopathology*, vol. 7, pp. 171-182.
- Horgan, T & Potrč, M 2010, 'The epistemic relevance of morphological content', *Acta Analytica*, vol. 25, pp. 155-173.
- Hurlburt, RT, Heavey, CL & Kelsey, JM 2013, 'Toward a phenomenology of inner speaking', *Consciousness and Cognition*, vol. 22, pp. 1477-1494.
- Inouye, T & Shimizu, A 1970, 'The electromyographic study of verbal hallucination', *Journal of Mental and Nervous Disease*, vol. 151, no. 6, pp. 415-422.
- Jackson, JH 1958, *Selected writings*, vol. 1, J Taylor (ed.), Basic Books, New York.
- Jaynes, J 1976, *The origins of consciousness in the breakdown of the bicameral mind*, Houghton Mifflin, Boston.
- Jeannerod, M 2006, *Motor cognition: what actions tell the self*, OUP, Oxford.
- Johns, LC, Nazroo, JY, Bebbington, P & Kuipers, E 2002, 'Occurrence of hallucinatory experiences in a community sample and ethnic variations', *British Journal of Psychiatry*, vol. 180, pp. 174-178.
- Jones, SR & Fernyhough, C 2007, 'Thought as action: inner speech, self-monitoring, and auditory verbal hallucinations', *Consciousness and Cognition*, vol. 16, pp. 391-399.
- Jorba, V & Vicente, A 2014, 'Cognitive phenomenology, access to contents, and inner speech', *Journal of Consciousness Studies*, vol. 21, no. 9-10, pp. 74-99.
- Kaplan, D 1990, 'Words', *Proceedings of the Aristotelian Society, Supplementary Volumes*, vol. 64, pp. 93-119.
- Kasher, A 1976, 'Conversational maxims and rationality', in A Kasher (ed.), *Language in focus: foundations, methods and systems: essays in Memory of Yehoshua Bar-Hillel*, D. Reidel Publishing Company, Dordrecht, pp. 197-215.
- Kasher, A 1982, 'Gricean inference revisited', *Philosophica*, vol. 29, no. 1, pp. 25-44.
- Keren, A 2007, 'Epistemic authority, testimony and the transmission of knowledge', *Episteme*, vol. 4, pp. 368-381.
- Knappik, F forthcoming, 'Bayes and the first person: consciousness of thoughts, inner speech and probabilistic inference', *Synthese*, doi: 10.1007/s11229-017-1321-3.
- Kohlberg L, Yaeger, J & Hjertholm, E 1968, 'Private speech: four studies and a review of theories', *Child Development*, vo. 39, pp. 691-736.
- Kriegel, U 2011, 'Cognitive phenomenology as the basis of unconscious content', in T Bayne & M Montague (eds.), *Cognitive phenomenology*, OUP, Oxford, pp. 79-102.
- Langland-Hassan, P 2008, 'Fractured phenomenologies: thought insertion, inner speech, and the puzzle of extraneity', *Mind & Language*, vol. 23, pp. 369-401.
- Langland-Hassan, P 2014, 'Inner speech and metacognition: in search of a connection', *Mind & Language*, vol. 29, no. 5, pp. 511-533.
- Larøi, F 2006, 'The phenomenological diversity of hallucinations: some theoretical and clinical implications', *Psychologica Belgica*, vol. 46, no. 1/2, pp. 163-183.
- Larøi, F, Sommer, IE, Blom, JD, Fernyhough, C, ffytche, DH, Hugdahl, K, Johns, LC, McCarthy-Jones, S, Preti, A, Raballo, A, Slotema, CW, Stephane, M & Waters, F 2012, 'The characteristic features of auditory verbal hallucinations in clinical and nonclinical groups:

- state-of-the-art overview and future directions', *Schizophrenia Bulletin*, vol. 38, no. 4, pp. 724-733.
- Lemaître, A 1905, 'Observations sur le langage interieur des enfants', *Archives de psychologie*, no. 4, pp. 1-43.
- Leudar, I, Thomas, P, McNally, D, & Glinski, A 1997, 'What voices can do with words: pragmatics of verbal hallucinations', *Psychological Medicine*, vol. 27, pp. 885-898.
- Lodge, D 1990, *After Bakhtin: essays on fiction and criticism*, Routledge, London.
- Machery, E 2005, 'You don't know how you think: introspection and language of thought', *British Journal for the Philosophy of Science*, vol. 56, pp. 469-485.
- MacKay, DG 1992, 'Constraints on theories of inner speech', in D Reisberg (ed.), *Auditory Imagery*, Erlbaum, Hillsdale, NJ, pp. 121-149.
- Malenka, RC, Angel, RW, Hamptom, B & Berger, PA 1982, 'Impaired central error correcting behaviour in schizophrenia', *Archives of General Psychiatry*, vol. 39, pp. 101-107.
- Mangan, B 2001, 'Sensation's ghost: the non-sensory "fringe" of consciousness', *Psyche*, vol. 7, no. 18.
- Martínez-Manrique, F & Vicente, A 2010, "'What the...!'" The role of inner speech in conscious thought', *Journal of Consciousness Studies*, vol. 17, no. 9-10, pp. 141-167.
- Martinich, AP 1980, 'Conversational maxims and some philosophical problems', *Philosophical Quarterly*, vol. 30, no. 120, pp. 215-228.
- McCarthy-Jones, S & Fernyhough, C 2011, 'The varieties of inner speech: links between quality of inner speech and psychopathological variables in a sample of young adults', *Consciousness and Cognition*, vol. 20, pp. 1586-1593.
- McCarthy-Jones, S, Trauer, T, Mackinnon, A, Sims, E, Thomas, N & Copoliv, DL 2014, 'A new phenomenological survey of auditory hallucinations: evidence for subtypes and implications for theory and practice', *Schizophrenia Bulletin*, vol. 40, no. 1, pp. 225-235.
- McGinn, C. (2009) Imagination, in Ansgar Beckermann, B. P. McLaughlin & S. Walter (eds), *The Oxford handbook of philosophy of mind*, Oxford: Oxford University Press.
- McGuigan, FJ 1966, 'Covert oral behaviour and auditory hallucinations', *Psychophysiology*, vol. 3, pp. 73-80.
- Megaw, ED 1972, 'Directional errors and their correction in a discrete tracking task', *Ergonomics*, vol. 15, pp. 633-643.
- Mele, AR 2001, *Self-deception unmasked*, Princeton University Press, Princeton, NJ.
- Mellor, DH 1993, 'Nothing like experience', *Proceedings of the Aristotelian Society*, vol. 93, pp. 1-16.
- Moore, GE 1942, 'A reply to my critics', in PA Schilp (ed.), *The philosophy of G. E. Moore*, George Banta Publishing Company, Menasha, WI.
- Moore, GE 1953, 'Propositions' in GE Moore, *Some main problems of philosophy*, George Allen & Unwin, London.
- Moseley, P & Wilkinson, S 2014, 'Inner speech is not so simple: a commentary on Cho & Wu (2013)', *Frontiers in Psychiatry*, vol. 5, no. 42, pp. 1-2.
- Nayani, TH & David, AS 1996, 'The auditory hallucination: a phenomenological survey', *Psychological Medicine*, vol. 26, pp. 177-189.

- Nichols, S 2003, 'Imagination and the puzzles of iteration', *Analysis*, vol. 63, no. 3, pp. 182-187.
- Pellegrini, AD 1981, 'The development of preschoolers' private speech', *Journal of Pragmatics*, vol. 5, pp. 445-458.
- Piaget, J 1959, *The language and thought of the child*, Routledge & Kegan Paul, London.
- Peirce, CS 1933, 'Existential graphs', in C Hartshorne & P Weiss (eds.), *Collected Papers of Charles Sanders Peirce, Volume IV: The Simplest Mathematics*, Harvard University Press, Cambridge, MA, pp. 291-470.
- Pinker, S 1994, *The language instinct*, William Morrow and Company, New York.
- Pitt, D 2004, 'The phenomenology of cognition or *What is it like to think that P?*', *Philosophy and Phenomenological Research*, vol. 69, no. 1, pp. 1-36.
- Pitt, D 2011, 'Introspection, phenomenality and the availability of intentional content', in T Bayne & M Montague, *Cognitive Phenomenology*, OUP, Oxford, pp. 141-173.
- Pizziconi, B 2006 'Politeness', in K Brown (ed.), *Encyclopedia of Language & Linguistics*, vol. 9, 2<sup>nd</sup> ed., Elsevier, Oxford, pp. 679-684.
- Proust, J 2006, 'Agency in schizophrenia from a control theory viewpoint', in N Sebanz & W Prinz (eds.), *Disorders of volition*, MIT Press, Cambridge, MA, pp. 87-118.
- Rabbitt, PMA 1966, 'Error-correction time without external signals', *Nature*, vol. 212, p. 438.
- Rieber, RW & Carton, AS 1987, 'Notes to references', in LS Vygotsky, *The collected works of L. S. Vygotsky*, vol. 2, JE Knox & CB Stevens (trans.), RW Rieber & AS Carton (eds.), Plenum Press, New York.
- Roessler, J 2013, 'Thought insertion, self-awareness, and rationality', in KWM Fulford, M Davies, RGT Gipps, G Graham, JZ Sadler, G Stanghellini & T Thornton (eds.) 2013, *The Oxford handbook of philosophy and psychiatry*, OUP, Oxford, pp. 658-672.
- Roessler, J 2016, 'Thinking, inner speech, and self-awareness', *Review of Philosophy and Psychology*, vol. 7, pp. 541-557.
- Romme, MA, Honig, A, Noorthoorn, EO & Escher, AD 1992, 'Coping with hearing voices: an emancipatory approach', *British Journal of Psychiatry*, vol. 161, pp. 99-103.
- Ross, A 1986, 'Why do we believe what we are told?' *Ratio*, vol. XXVII, no. 1, pp. 69-88.
- Ryle, G 1949, *The concept of mind*, Hutchinson, London.
- Sass, LA 1994, *The Paradoxes of Delusion: Wittgenstein, Schreber, and the schizophrenic mind*, Cornell University Press, Ithaca, NY.
- Seal, M, Aleman, A & McGuire, P 2004, 'Compelling imagery, unanticipated speech and deceptive memory: neurocognitive models of auditory verbal hallucinations in schizophrenia', *Cognitive Neuropsychiatry*, vol. 9, no. 1/2, pp. 43-72.
- Searle, J 1969, *Speech acts: an essay in the philosophy of language*, CUP, Cambridge.
- Sedman, G 1966, "'Inner voices": phenomenological and clinical aspects', *British Journal of Psychiatry*, vol. 112, pp. 485-490.
- Shergill, SS, Bullmore, ET, Brammer, MJ, Williams, SCR, Murray, RM & McGuire, PK 2001, 'A functional study of auditory verbal imagery', *Psychological Medicine*, vol. 31, no. 2, pp. 241-253.

- Siewert, C 1998, *The significance of consciousness*, Princeton University Press, Princeton, NJ.
- Siewert, C 2011, 'Phenomenal thought', in T Bayne & M Montague (eds.), *Cognitive phenomenology*, OUP, Oxford, pp. 236-267.
- Smithies, D 2013, 'The nature of cognitive phenomenology', *Philosophy Compass*, vol. 8, no. 8, pp. 744-754.
- Sommer, IEC, Daalman, K, Rietkerk, T, Diederer, KM, Bakker, S, Wijkstra, J & Boks, MPM 2010, 'Healthy individuals with auditory verbal hallucinations; who are they? Psychiatric assessments of a selected sample of 103 subjects', *Schizophrenia Bulletin*, vol. 36, no. 3, pp. 633-641.
- Sperber, D & Wilson, D 1995, *Relevance: communication and cognition*, 2<sup>nd</sup> ed., Blackwell, Oxford.
- Stephane, M, Barton, S & Boutros, NN 2001, 'Auditory verbal hallucinations and dysfunction of the neural substrates of speech', *Schizophrenia Research*, vol. 50, pp. 61-78.
- Stephane, M, Thuras, P, Nasrallah, H & Georgopoulos, AP 2003, 'The internal structure of the phenomenology of auditory hallucinations', *Schizophrenia Research*, vol. 61, pp. 185-193.
- Strawson, G 1994, *Mental reality*, MIT Press, Cambridge, MA.
- Strawson, G 2003, 'Mental ballistics or the involuntariness of spontaneity', *Proceedings of the Aristotelian Society*, vol. 103, pp. 227-256.
- Tappan, MB 1997, 'Language, culture, and moral development: a Vygotskian perspective', *Developmental Review*, vol. 17, pp. 78-100.
- Tye, M & Wright, B 2011, 'Is there a phenomenology of thought?', in T Bayne & M Montague (eds.), *Cognitive phenomenology*, OUP, Oxford, pp. 326-344.
- van der Veer, R & Yasnitsky A 2011, 'Vygotsky in English: what still needs to be done', *Integrative Psychological & Behavioral Science*, vol. 45, pp. 475-493.
- Vendler, Z 1979, 'Vicarious experience', *Revue de Métaphysique et de Morale*, vol. 84, no. 2, pp. 161-173.
- Vicente, A & Martínez-Manrique, M 2011, 'Inner speech: nature and functions', *Philosophy Compass*, vol. 6, no. 3, pp. 209-219.
- Voloshinov, VN 1973, *Marxism and the philosophy of language*, Seminar Press, New York.
- Vygotsky, LS 1931/1997, 'Genesis of higher mental functions', in RW Rieber (ed.), *The collected works of L. S. Vygotsky*, vol. 4, Plenum Press, New York.
- Vygotsky, LS 1934/1986, *Thought and language*, A Kozulin (trans. and ed.), MIT Press, Cambridge, MA.
- Vygotsky, LS 1934/1987, 'Thinking and speech', N Minick (trans.), in RW Rieber & AS Carton (eds.), *The collected works of L. S. Vygotsky*, vol. 1, Plenum, New York, pp. 39-285.
- Wegner, DM 2002, *The illusion of conscious will*, MIT Press, Cambridge, MA.
- Weiskrantz, L, Elliot, J & Darlington, C 1971, 'Preliminary observations of tickling oneself', *Nature*, vol. 230, pp. 598-99.
- Wertsch, JV 1979, 'The regulation of human action and the given-new organization of private speech', in G Zivin (ed.), *The development of self-regulation through private speech*, Wiley, New York, pp. 79-98.

Wertsch, JV 1980, 'The significance of dialogue in Vygotsky's account of social, egocentric, and inner speech', *Contemporary Educational Psychology*, vol. 5, pp. 150-162.

Wertsch, JV 1991, *Voices of the mind: a sociocultural approach to mediated action*, Harvester Wheatsheaf, Hemel Hempstead.

Wertsch, JV & Stone, CA 1985, 'The concept of internalization in Vygotsky's account of the genesis of higher mental functions', in JV Wertsch (ed.), *Culture, communication and cognition: Vygotskian perspectives*, CUP, Cambridge, pp. 162-179.

Wetzel, L 2014, 'Types and tokens', in EN Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, (Spring 2014 Edition) viewed 8 May 2017  
<<https://plato.stanford.edu/archives/spr2014/entries/types-tokens/>>.

Wiley, N 2016, *Inner speech and the dialogical self*, Temple University Press, Philadelphia.

Williams, B 1973, 'Imagination and the Self', in B Williams, *Problems of the self: philosophical papers 1956-1972*, CUP, Cambridge, pp. 26-45.

Winsler, A, De León, JR, Wallace, BA, Carlton, MP & Willson-Quayle, A 2003, 'Private speech in preschool children: developmental stability and change, across-task consistency, and relations with classroom behaviour', *Journal of Child Language*, vol. 30, pp. 583-608.

Wittgenstein, L 1953/2001, *Philosophical Investigations*, GEM Anscombe (trans.), Blackwell, Malden, MA.

Wu, W 2012, 'Explaining schizophrenia: auditory verbal hallucination and self-monitoring', *Mind & Language*, vol. 27, no. 1, pp. 86-107.