

Andreas Kemmerling  
(Universität Heidelberg)

*First Person Authority  
without Glamorous Self-Knowledge*

In this talk, I want to discredit the view that we have a very special sort of knowledge about our own current mental states, events or processes. Of course, I do not deny that we have self-knowledge, but I do want to cast serious doubt on the idea that some of our self-knowledge is far out of the ordinary. Today, I shall focus exclusively on the case of our current conscious thoughts. This case would be a paradigmatic example of the special self-knowledge in question, if there were such a thing.

I want to attack what I take to be the best reason for believing in extraordinary self-knowledge – the assumption that it is needed for explaining the special authority we have when it comes to *saying* what our current mental phenomena are – “first person authority”, as it is called. I accept without reservation that we do have it. But although I accept special *authority*, I find it hard to believe in a very special sort of *knowledge* – knowledge of contingent matters more profound than any other. I shall not attack here the general thesis that there is such knowledge. What I shall try to refute is the claim that it is only in virtue of such special knowledge that first person authority can be explained.

For this purpose I shall sketch a different kind of explanation which does not presuppose self-knowledge and which is meant to show this: All that is needed, in order to say (with authority) what one thinks, is an ability every normal speaker has – namely, the ability to think out loud. Or so I shall argue. A spin-off of my way of explaining first person authority is that it reveals why it is so strange to speak of knowledge at all in this context. Let me anticipate this one point. It is

this: first person authority bears on *truth guaranteed by honesty*. But if honesty is sufficient for truth-that-cannot-be-doubted, then knowledge is uncalled-for.

### 1. The target: Glamorous self-knowledge of sayable thoughts

We have self-knowledge of various sorts: knowledge of things we have done or suffered, for example, and some knowledge of who we are: of our character-traits, our temper, our inclinations, weaknesses, feelings, addictions, worries, lusts and so on. Most of this knowledge is human knowledge of the regular kind, nothing exciting about it, epistemologically speaking. Even where we actually do know, we could have been mistaken.

But many philosophers assume that some of our self-knowledge is of a very special kind: it is such that we cannot be mistaken about it. It is *infallible* self-knowledge in the following sense: If it seems to a subject that he so-and-sos (or if he believes that he so-and-sos), then it is somehow guaranteed that he actually so-and-sos. Our current conscious thoughts are one striking example. If it seems to a human subject  $S$ , at time  $t$ , that he is entertaining a particular thought (*e.g.*, the thought he would express by the words “Harvey is stupid”), then, in this exact moment, he incontestably entertains this thought. Another possibly more controversial example would be one’s current beliefs: If it seems to  $S$  (or if he believes) that he holds the belief which he would express by assenting to “Harvey is stupid”, then, incontestably, he does hold this belief, in the exact moment in which it seems so to him. As I said, I shall focus today on the case of current thoughts, and leave the issue of our alleged self-knowledge concerning our beliefs for another occasion.

Let me warn you about the slightly technical way in which I shall use the phrase “to think that ...”. In saying that  $S$  is thinking that such-and-such, I do not mean to suggest that  $S$  judges, or is

at least inclined to judge, that such-and-such. What I have in mind is some broader (and, possibly, artificial) sense of *thinking that*, namely: having or entertaining, in whatever mode, the thought in question. (Modes of entertaining a thought are, for example: judging, doubting, possibly merely “perceiving” it in Descartes’s sense, or “grasping” it in Frege’s sense; but there are many more modes, and I shall briefly return to this point later on). One more technicality. When I say that *S* thinks that such-and-such, the correctness of what is inserted for the placeholder phrase “such-and-such” is determined by the way *S* himself would express the thought he is entertaining, if he were a normal competent speaker of English who expresses his thought in English. The reason for this is to avoid side-issues which otherwise would have to be considered very carefully (for example issues about the compatibility of anti-individualism and infallible self-knowledge). – In brief, for the purposes of this paper, whenever I say something like “*S* (at time *t*) thinks that Harvey is stupid”, please understand this as an abbreviation for:

At *t*, *S* is entertaining, in whatever mode, a thought which he would express at *t*, if he were to express it, at *t*, in English as a normal competent speaker of this language, by using the very words “Harvey is stupid”.<sup>1</sup>

Many philosophers have assumed that we have knowledge about what we are currently thinking (in the sense of “think” I’ve just explained). This kind of alleged self-knowledge has been showered with highly honorific epithets like *direct*, *a priori*, *infallible*, *self-evident*, *incorrigible*, *self-verifying*, *immediate*, *arrived at via privileged access*, *resulting from self-intimation*, etc. And it is widely assumed that in voicing one’s current thoughts, one has so-called first person authority.<sup>2</sup> Moreover, this is *substantial* (non-trivial) self-knowledge. By this I mean that the truth of what is known is contingent and is not guaranteed solely by one’s sheer existence or one’s essential properties (like the knowledge of one’s being self-identical, or one’s having a mind). If somebody were to know that he is thinking that Harvey is stupid, then he would know a substantial fact; for

he might as well have had another thought instead at this particular moment. – Self-knowledge, which deserved these characterizations (or at least sufficiently many of them), would be what I shall call (*epistemologically*) *glamorous*, as opposed to the more standard sorts of knowledge which we have about ourselves.

I have certain misgivings about the applicability of some of the attributes I've just mentioned, and about the appropriateness of speaking, in this context, of knowledge at all. After all, to say such a thing as “Birgit knows that she is thinking that Harvey is stupid” sounds somewhat strange. To exercise a certain amount of conceptual caution here, is, as far as I can tell, not just a silly quirk. You don't have to have been under the influence of philosophers like Wittgenstein, Austin or Ryle, to feel that something may be conceptually amiss with such ways of talking. Even Descartes, assumed by many to have been a grand champion of self-knowledge about our current thoughts, would not have put it this way; he would have preferred to say that Birgit is conscious of her thought/thinking, or that she ‘experiences’ that she thinks. (What you find in Descartes' writings is not: *scio me cogitare*, but instead phrases like: *meae cogitationis conscius sum* or *experior me cogitare*.<sup>3</sup>) According to Kant's explication of the concept of knowledge in his *Kritik der reinen Vernunft* (1787, A 822/B 850), we have no knowledge about our own current mental states. For knowledge entails “objective certainty” and that is, as he puts it, certainty for everybody. Wittgenstein's opposition to speaking of knowledge in this context is public knowledge. And, to return to the current discussion, Fred Dretske argues that there is no knowledge, and certainly no glamorous knowledge, about the fact *that* one is thinking, even if it were taken for granted that one knows *what* one thinks.<sup>4</sup> The reasons Dretske puts forward for his, as he labels it, “skepticism” about self-knowledge are clearly different from those which might be found in Descartes, Kant or Wittgenstein. So conceptual caution as to the appropriateness of speaking of knowledge at all, in this context, should not be simply shrugged off. Such caution does not betray undue fondness for being linguistically picky. First rate thinkers, past and present, with quite

different outlooks on philosophy and the human mind, have given us reason to take this issue seriously. – So we should not take it for granted that we have glamorous knowledge, or even only regular knowledge, of our current thinking. But to question this is, of course, not to question that we are aware of thinking the thoughts we are currently thinking. Rather, it is to doubt that the awareness, or consciousness, we have of our current thinking is a sort of knowledge.

\*

As a starting point for the following discussion let us assume the following conditional:

If there is glamorous self-knowledge, then some of our current thoughts are items about which we can have it.

Which of our current thoughts? Well, at least those which I shall label as *sayables*, *i.e.*, thoughts which one may formulate in words, completely and exactly as one has them. I assume that some of our conscious thoughts are sayables. And I intend this to be a modest thesis. I am not making any grand claim here about the nature of thinking. I am not saying, like Kant, that “thinking is talking with oneself ... and therefore it is also hearing oneself inwardly” (1800, Teil I, § 36). Neither do I want to claim that we, at least sometimes, ‘think in language’.

Let me add that having a sayable thought is not the same as merely having a fully meaningful linguistic item going through one’s mind – even if one fully understands the item. A while ago, I couldn’t get rid of a certain line from a song by an English pop group; sadly enough, the words “Reason is treason” lingered in my mind for weeks. But it happened only very rarely during this period that I ever *thought*, even if only in the sense explained, that reason is treason. – I have no

theory to offer about sayable thoughts, but I shall assume that they exist and that we could have glamorous self-knowledge about some of them, if we could have such knowledge at all.

To sum up these preliminary remarks: If there is glamorous self-knowledge (in the following often abbreviated as *GSK*), it is knowledge about a contingent fact about oneself. It is substantial, *i.e.*, cannot be derived from knowing that one exists and has such-&-such essential properties. And it is at least infallible in the sense explained (maybe it has further epistemological qualities like self-evidence etc.) The friend of *GSK* holds that there is such knowledge and that we have it, *inter alia*, about our current conscious sayable thoughts. I doubt that there is *GSK*. But I shall take issue only with a certain claim about it, namely that it is needed in order to explain first person authority.

## 2. Supreme authoritativeness in saying what one thinks

Now let's turn to our ability to *say* what we think. Let me first mention an ambiguity of the phrase "(being able) to say what one thinks". In one reading the subordinate *wh*-clause is an indirect interrogative and in another it is a relative clause. There are languages, like English or German, which blur this difference by not marking it syntactically. But there are also languages, like Latin, in which the difference is marked syntactically. In Latin, the indirect-interrogative reading of the *wh*-clause of "Harvey says what he thinks" is given by a "quid"-clause in the subjunctive mood ("Herveus dicit *quid* se ipse cogitat"), the relative-clause reading by a "quod"-clause in the indicative mood ("Herveus dicit *quod* se ipse cogitat"). For the sake of brevity, I shall refer to the two readings as the *quid*-reading and the *quod*-reading, respectively.<sup>5</sup>

The difference between these readings is quite manifest in sentences like "Harvey knows what he thinks". If this sentence, in the *quod*-reading, is taken in combination with "What Harvey thinks is that it's raining", it follows that Harvey knows that it's raining. This consequence cannot be

drawn if the first sentence is taken in its *quid*-reading. (Interestingly, “to believe what one thinks” seems to have no *quid*-reading.) – The ambiguity may be less perspicuous in “being able to say what one thinks”. But it is there nevertheless. In ascribing to *S* the ability to say what he thinks, we may commit ourselves either to the claim:

(Quod) If *S* thinks that ---, then he is able to say “...” (where his uttering “...” is a way of exactly saying that ---);

or instead, more naturally, to the claim

(Quid) If *S* thinks that ---, then he is able to say that (or whether) he thinks that ---.

In this paper, I shall be primarily concerned with the *quid*-sense of our ability to say what we think. But I shall argue that, very roughly speaking, the ability mentioned in (Quod) is explanatory of the ability (Quid) is about.

The way in which we can say what we currently think has some epistemological splendour too. At least if the thought in question is a sayable one. Consider a case in which you say what you think by using a sentence of the type “I’m thinking that such-&-such”. What you thereby say is that you are thinking that such-&-such. Assume that the thought that such-&-such is one of your sayable thoughts. Now in such a case (a) you yourself, given that you speak honestly, cannot doubt what you say, (b) nobody else, except those who have reservations about your honesty, can reasonably doubt what you say<sup>6</sup>, and (c) nobody else but you could say, in such a doubt-excluding way, what *you* think. Let us say that

*S*, in a certain context *c*, says that *p* **with supreme authoritativeness**, if the following holds: Whoever assumes that *S* is honest in his saying that *p* (in *c*), cannot rationally doubt

that  $p$ ; but if somebody other than  $S$  were to honestly say the same,<sup>7</sup> it could be coherently doubted that  $p$ .

This sort of authority is marked by a distinctive exclusiveness: what one person can say with supreme authoritativeness, only he or she can say in this way. Put differently: It is marked by a distinctive asymmetry, between him, who says something with this sort of authority, and everybody else. His honesty is sufficient for the indubitability of what he says; whereas honesty would not be sufficient, if anybody else said it. I think that my characterization of supreme authoritativeness captures what is meant by the more common term “first person authority”. So I shall use these two terms interchangeably.

What can be said with this special authority? Well, clearly anything we can say by using a sentence of the type “I am thinking that  $p$ ” as a report about oneself and one’s current thinking. Such a ‘real-time autobiographical’ use may be rare or even somewhat *recherché*. But I accept that it exists and is of some philosophical interest. There may be other things which we can say with this special authority (“I’m in pain”, “I believe ...”, “I feel ...”, etc.). But, as I said, I shall focus on first person *thought* ascriptions. Here is an example of what I shall consider as an uncontroversial starting point for my ensuing considerations: Given that the thought that someone would express by saying “Harvey is stupid” is one of her sayable thoughts, then what she would say in using the words “I am thinking that Harvey is stupid” is something she would say with supreme authoritativeness.

My central concern in the rest of this talk will be the question: How is this authority to be explained? Or, more specifically: What could plausibly be offered as an explanation, if we abstain from the usual one (“We have this ability because we have self-knowledge”)?

I shall restrict myself in this talk to cases in which the subject is a normal speaker of a natural language,<sup>8</sup> where the thought in question is sayable and where the overall circumstances of the utterance are normal. I think it's a good idea to concentrate on such cases, because it is primarily within these limits that we should have some confidence in our judgments concerning the matters at issue.

Given these restrictions, there is an intimate connection between alleged *GSK* about our current sayable thoughts on the one hand, and the ability to say authoritatively what one thinks, on the other. In fact, I assume that the statements (1) and (2) below are conceptual truths.

Let *S* be a normal speaker of a natural language, who (at time *t*) has a sayable thought that *p*.

- (1) If *S* at *t* has *GSK* about (his having) the thought in question, then, *ceteris paribus*, *S* at *t* has the ability to say authoritatively that he, at this moment, is thinking that *p*.
- (2) If *S* at *t* has the ability to answer authoritatively the question whether he, at this moment, is thinking that *p*, then: If there is *GSK* at all, then, *ceteris paribus*, *S* at *t* has *GSK* about (his having) this thought.

Notice that (1) and (2) contain a *ceteris paribus* clause. This is crucial for their being conceptual truths. Without this qualification, they wouldn't even be true. Statements that are conceptually true if they contain such a qualification, but may well be false if they don't, I call *cp-necessities*.<sup>9</sup> – Also notice that the existence of *GSK* is presupposed neither by (1) nor (2). Both can be readily accepted even by someone, like me, who doesn't believe in *GSK*. – I consider these statements as common ground between the friends of *GSK* and me.

\*

But given such a close conceptual connection between first person authority and *GSK*, how can one believe in the one but not in the other? Isn't the existence of *GSK* suggested by an inference to the best (in fact, the only available) explanation? "You say you don't deny first person authority. But how on earth could this authority be explained – except by the fact that we have it in virtue of our *GSK* about our current thoughts?" – This is a fair question which I shall try to answer.

For this purpose, I shall introduce an ability we all have, describe some of its characteristic peculiarities, and try to show that it opens a way to account for the authority with which we can say what our current sayable thoughts are. What makes this sort of explanation interesting is that it does not presuppose self-knowledge, glamorous or not.

Later on, in section 5 below, I shall argue that in acting on this ability we may indeed acquire self-knowledge properly so-called, but that the knowledge acquired in this way is regular knowledge: neither infallible nor otherwise glamorous.

### 3. The ability to think out loud

*No, I must not try to think, simply utter.*  
Samuel Beckett

Now what explains our ability to say, authoritatively, what we think? I suggest that it is our ability to think out loud. By this I mean the ability to *'put out'*, *manifest*, or *display* the thought one is consciously entertaining, in whatever mode, in this exact moment – *i.e.*, the thought one is having as one puts it out.

Thoughts can be entertained in various modes, many more than those which are usually considered in the philosophical literature. We can entertain a thought curiously ("what does it

amount to?”), critically (“what are its weaknesses?”), affirmatively (as accepted premises in a process of consideration), playfully (as a toy, as it were), daydreamingly, fictionally (as something which may go into the movie-script one is writing), hedonistically (as something which to think is delightfully pleasant), etc., etc.

In thinking out loud, we put out the very thought we are having.<sup>10</sup> The question *in which mode* the thought is entertained is not settled (and not even addressed) by what we do when we think out loud. Metaphorically speaking, in thinking out loud, we turn the mute function off and present sayable thoughts of our current thinking in their original version (without subtitles of any kind).

It should be noted that if and when we act on this ability in public –which, by the way, is something we do only very rarely, fortunately–, we do not perform *constative* or *assertive* speech acts. In fact, there need no commitment at all concerning the *truth* of what we utter. We say what we think, in the *quod*-sense of verbally exposing that which we think. But this does not entail that the thought we are thinking is one we hold true, nor does it mean that by our utterance we report, or assert, that we are thinking the thought we put out. Moreover, the honesty (or frankness) which it characteristically takes a sane adult human being to think out loud, at least in the presence of an audience, is one of its *conceptual* ingredients. Someone who acts as if he were thinking out loud but in fact tries to hide several of the sayable thoughts he is consciously entertaining, is not really engaged in what he pretends to be doing. He would not be thinking out loud; rather he’d be acting dishonestly. Note also that the honesty involved in thinking out loud does not concern one’s believing to be true what one says, but rather one’s voicing what one is thinking – even if the thought voiced is not one which one holds to be true. So this sort of honesty has more to do with unreserved undisguisedness than with ‘honest *truth*’. Taken to extremes it amounts to mental exhibitionism, not to unyielding veracity.<sup>11</sup>

In pure acts of thinking out loud<sup>12</sup>, it might be argued, we do not even perform any communicative speech acts. At least, we clearly do not, then, mean anything by our utterances, in the Gricean sense of “to mean something by one’s utterance”. The famous Gricean mechanism is not at work when we think out loud. We don’t intend, when we’re thinking out loud, to get our audience to recognize our higher order intention to recognize our basic intention to bring about, in them, particular beliefs. In fact, we do not even intend to produce beliefs in our audience in the first place, neither beliefs about the topics of our thoughts on display (*e.g.*, Harvey’s being stupid or not) nor beliefs about our own beliefs concerning these topics.

The speech act of thinking out loud seems not to fit in the usual taxonomies which focus on communicative speech acts. Superficially, it may seem to bear some similarity with what Searle has called *declarations*: speech acts like declaring war or making a gift by uttering sentences like “It is war” or “This is yours”.<sup>13</sup> But clearly the similarity breaks down at a crucial point: the performance of a declarative speech act brings into existence the very state of affairs which corresponds to the propositional content of the speech act. But when someone says, in the course of thinking out loud, “Harvey is stupid”, he thereby neither brings it about that Harvey is stupid, nor does he thereby bring about his own thought that Harvey is stupid. Furthermore, a thinking out loud is not an *expressive* like thanking, apologizing, congratulating etc. In performing an expressive the speaker presupposes the truth of the expressed proposition; but it is an essential characteristic of a thinking out loud that it has no such presupposition.

A speaker actually expresses some of his attitudes in performing a common speech act (*e.g.*, an assertive, commissive, or directive) which is not formally an expressive. That is to say, he represents himself as having certain beliefs, desires, intentions, preferences, etc. Whereas in a pure thinking out loud nothing is *re-presented* as being the case. Moreover, it would be misleading to say that the speaker expresses that he is thinking the very thought which he thinks

out loud. For to express an attitude (in the sense which is at issue here) is to commit oneself to having the attitude one represents oneself to have. No such commitment is involved in a thinking out loud. – Maybe the term *Exhibitive* would be apt for labelling the category of speech acts to which thinking out loud belongs. It is the characteristic point of such speech acts that the speaker, in making his utterance, exhibits (presents, rather than represents) himself as being so-and-so.

The speaker, in thinking out loud, does not express (in the relevant sense) that he is thinking this-&-that. His utterance is a manifestation – like putting the money in cash on the table, which manifests your having it, as opposed to signing a cheque. One's entertaining the thought that Harvey is stupid, or rather one's entertaining it, can be expressed (or attested to) in many ways, for example by raising one's arm in reaction to the request "Who of you is, right now, entertaining the thought that Harvey is stupid? Please raise your arm exactly when you are entertaining this thought, in whatever mode!". But in doing so, the arm-raising person does not *manifest* his having this thought. Even if he is honest, there is leeway for doubt: he may have misheard the request, his arm-raising may be due to a sudden cramp, and so on. A manifestation, *sensu stricto*, of *X* is something that leaves no doubt at all about the existence of *X*. If honesty is granted, thinking out loud counts as a manifestation of the entertaining of the very thought which is displayed acoustically. In normal circumstances, further evidence is neither needed nor would it be clear what it could consist in. Possibly, thinking out loud (with honesty granted) is the only way to manifest, in the realm of the sayable, that we are thinking the particular thought which we think out loud. It's imaginable that people who are very quick scribblers could do the same in writing. (Remember the automatic writing movement among the Dadaists and some of the Surrealists.) Maybe someone like Jackson Pollock was able to do something of this kind with oil on canvas<sup>14</sup>; someone like Charlie Parker with a saxophone, someone like Sonny Sharrock

with a guitar, etc. – I’m not sure about it, but it seems to me that some of our non-sayable thoughts can also be manifested, in some way analogous to thinking out loud.

We don’t do a lot of (pure) thinking out loud, at least not in the presence of an audience. But whether we ever do it or not, we are able to do it. Any normal speaker has this ability. And when we do it, we are often not aware of doing it. A friend once told me he never does it and more than that: he simply *couldn’t* do it; he couldn’t even try. A little later when we were playing chess, I heard him mumbling things like: “I move it there, then he will capture it. But that’s O.K. Then I can threaten his queen. Ahem. But that’s no good. Ahem. No, this won’t work. Once again, I move it there, ...” Sound volume is not important. Thinking out *sotto voce* is thinking out loud, albeit in a low voice.

#### 4. The ability to think out loud explanatory of the ability to say what one thinks

Now here is one crucial fact about the ability to think out loud, in virtue of which it is explanatory of the ability to say, with absolute authority, what one thinks: Given overall normal circumstances, it is conceptually sufficient. This is to say: (3) is a conceptual truth.

- (3) *Ceteris paribus*: If *S* is able to think out loud (in *L*), then he is able to say authoritatively (in *L*), what he thinks.

Why should this be a *cp*-necessity? And why should its antecedent be *explanatory* of the consequent?

As to the first question, and leaving the authority issue aside for a moment, it needs no argument to see that anyone who is able to think out loud is normally able to say what he thinks. If he has the ability to present, by uttering “So-&-so”, what he’s thinking –and is, by assumption, a normal

speaker and has therefore mastered the use of “I am thinking that ...”–, then, in overall normal circumstances, he is also able to say “I’m thinking that so-&-so”, thereby telling us what he he is thinking. – If you don’t think this is obvious, see if the following consideration persuades you: If *S* is able to think out loud, he is able to verbally present (to *quod*-say) what he thinks – *e.g.*, by uttering “Harvey is stupid”, thereby performing the exhibitivie speech act of thinking out loud. But for anyone who has mastered the self-ascriptive use of “I’m thinking that ...”, it follows from this (excuse me for being excessively banal) that he is also able to *quid*-say what he thinks – by uttering “I’m thinking that Harvey is stupid”, therein performing a truth-evaluable speech act of self-ascription. So (3) is plainly true, if the occurrence of “authoritatively” is left aside. Moreover, in order to recognize this plain truth we don’t have to engage in empirical observation or speculation; it is enough that we have mastered the relevant concepts (of a normal speaker, the ability to think out loud, of a situation in which overall normal conditions obtain, etc.).

Now to the authority, which I claim to be a *ϕ*-conceptual consequence of the ability to think out loud. Where is it supposed to come from? Well, basically from two things:

(4) *Ceteris paribus*, in saying what he thinks, the speaker *could*, if he wished, act on his ability to think out loud.

(5) *Ceteris paribus*, if in fact he acts on his ability to think out loud, his saying what he says has all the marks of supreme authoritativeness in saying what he thinks.

As to (4), I shall sketch my reasons for it in section 7 below. Let me at this point merely try to forestall one objection: Yes, there may be people who are psychologically inhibited, people who in overall normal circumstances could not bring themselves to think out loud (not even under

their breath), however hard they try. Yet if they are *normal* competent speakers, they still have the ability. What they'd lack would be the second order ability to implement it.

As to (5), let us first consider why the *ϕ*-proviso is required. In acting on the ability to think out loud, a speaker is not immune to the usual linguistic mishaps. He may fall victim to a slip of the tongue, momentary linguistic confusion, etc. In such a case, although acting on the ability, he would not say what he thinks. *A fortiori* he wouldn't authoritatively say what he thinks. This has to do with a general fact about abilities. Acting on an ability doesn't guarantee that the desired result is achieved. Even if I fail in my attempt to jump a metre-high fence, I may have nevertheless brought to bear, on that occasion, my ability to jump much higher than that: somebody may have pulled me down when I was high up in the air. Similarly, when in trying to say what I think, I act on my ability to think out loud, I may nevertheless fail to say what I think: I may misspeak. The ability to think out loud isn't magical. Something may intervene. Maybe not only slips of the tongue, but also other disruptive factors. Since there is no hope for a complete list of possibly relevant abnormalities (there may be ones we haven't even thought of), a *ϕ*-clause is needed to exclude them. The best we can safely say is this:

Given overall normal conditions, when a normal speaker, in acting on his ability to think out loud, says what he thinks with supreme authoritativeness.

And this is nothing but a reformulation of (5).

Let us focus now on cases in which we presume overall normal conditions. In such cases it holds that given the speaker's honesty in saying what he thinks, the truth of what he says cannot be doubted. For as long as we, the others, presume that he, in saying "I'm thinking that so-&-so", acts on his ability to think out loud, we can have no reason to doubt that he is thinking that so-

&-so. Moreover, in acting on his ability to think out loud, he simply cannot mistakenly give us something he actually does not think, in this exact moment, as something he's thinking. (Remember that we are not talking about what  $S$  holds true or judges.) There is just no scope for involuntary mistakes on his part which are not covered by the  $\varphi$ -clause. Errors of memory, for example, are out of the question because of the simultaneity. Such errors would not be out of the question, if instead he were to say: "One moment ago, I had exactly the same thought that you just mentioned: Harvey is stupid". Even if he's an honest person to the bone, he may err in *this*. His actual thought then may have been that Harvey is silly. But if in acting on his ability to think out loud, he says "I'm thinking that Harvey is stupid" (and the situation is normal, as we presume), he simply cannot be mistaken in some such way; he puts out the thought as and how he is entertaining it.

So if he acts on this ability, then (presuming overall normality) the question "But was it *really* what you were thinking at this particular moment?" has only one correct answer: Yes. – In a nutshell, deceitfulness is excluded by the speaker's honesty, involuntary mishaps are ruled out by the normality constraint. In the light of the exhibitiv nature of the act, these are the only sources for coherent doubt of the relevant kind. That is to say, there can be no coherent doubt about the truth of what *he* says (when he utters "I'm thinking that so-&-so"), thereby acting on his ability to think out loud. But if we said the same (*i.e.*, that he is thinking that so-&-so), the truth of what *we* say can be doubted. Therefore he speaks with supreme authoritativeness. Hence (5) is true.

\*

It may be asked: But how, then, could we ever *know* in a given case whether it was a thinking out loud? We would have to know that the speaker wasn't being dishonest. And even if we knew this somehow, how could we *know* that he didn't misfire – that the circumstances of his utterance

were in fact normal? These are fair questions. (And I tend to think that the answer is: We cannot know.) But if such questions were meant as an objection, they'd be off the mark. My claim is that we cannot coherently doubt that we are facing a case of thinking out loud, *if we presume* the absence of dishonesty and abnormalities. The question whether we can ever know that our presumptions are true is of no concern here.

But again, one may feel uneasy about the requirement, or stipulation, that honesty be a conceptual ingredient of thinking out loud. Two remarks on this. First, I hold that this is simply the right way of explicating the concept. Just as the appropriate analysis of informing includes truth (otherwise one has, at best, been misinformed),<sup>15</sup> the analysis of thinking out loud should require the speaker's honesty. If the speaker is not honest, if he, *e.g.*, recites sentences with his mind elsewhere, his utterance ought not to be described as a thinking out loud. Just as there are not two types of informing (truly and falsely), there are not two types of thinking out loud (honest and dishonest). Secondly, no vicious circle is involved in such an explication. The relevant honesty is not to be characterized as *saying what one thinks*, or suchlike. It is simply the absence of any pertinent intention to conceal, mask or withhold the sayable thought in question. Put in old fashioned terminology, it's not a 'real and positive' cognitive attitude but rather a 'privation': the absence of a certain family of such perfidious attitudes.

\*

So much in support of (4) and (5), which in turn support (3). Let's assume then that someone's being able to think out loud is conceptually sufficient<sup>16</sup> for his being able to say authoritatively what he thinks. The more ambitious claim is this: The fact that a speaker has the first ability *explains* that he has the second. What sort of explanation is this? Given the conceptual truth of (3), it's not a causal explanation. What then is it? A conceptual explanation.

What is this? Consider the sentence “*S* is able to say authoritatively what he thinks, *in virtue of* his being able to think out loud”. For it to be true, the truth of

- (3) *Ceteris paribus*: If *S* is able to think out loud (in *L*), then he is able to say authoritatively, (in *L*), what he thinks.

is not enough. Mere conceptual sufficiency is not explanatory. (Mr Bird is neither unmarried, nor male, nor adult *in virtue of* his being a bachelor.) It may be even denied that a conceptual relationship can be explanatory at all. Admittedly, in real life conceptual explanations are rare; usually they are not an appropriate reaction to the request for an explanation.<sup>17</sup> I shall not try to argue here that such relationships can be genuinely explanatory (in particular in philosophy), but rather assume that they are.<sup>18</sup>

Yet, even if the admissibility and potential value of conceptual explanations is granted in general, a question might be raised about the particular explanation I am proposing. It can be put as follows: “Might we not just as well do it the other way round, *i.e.*, ‘explain’ the ability to think out loud in terms of the ability to say what one is thinking? But if this works equally well (or equally poorly), then this shows that your alleged conceptual explanation is no genuine explanation. To pick up your own example: Mr Bird’s bachelorship can be explained in terms of his being adult, unmarried, etc. only because we *cannot* explain the fact that he is adult, unmarried etc. in terms of his being a bachelor. If there is as much reason for the truth of ‘*A* in virtue of *B*’ as there is for the truth of ‘*B* in virtue of *A*’, neither sentence expresses a genuine explanation”. – I think that this is correct. More than that, it is especially relevant for our considerations because the reversal of (3), namely

(R3) *Ceteris paribus*: If  $S$  is able to say authoritatively, (in  $L$ ), what he thinks, then he is able to think out loud (in  $L$ ),

is also a conceptual truth. That is to say that the two abilities at issue here are conceptually sufficient for one another – on the assumption that  $S$  is a normal speaker. (This qualification is crucial for the argument below.)

In order now to answer this point let me first of all specify the conditions which have to hold, in general, for an ability  $A_1$  (in regard to a certain group  $G$ ) to be conceptually explanatory of an ability  $A_2$ : namely, (i) for members of  $G$ , having  $A_1$  is conceptually sufficient for having  $A_2$ , (ii)  $A_1$  and  $A_2$  are distinct abilities, (iii)  $A_1$  can be acquired without acquiring  $A_2$ , but (iv) not *vice versa*.

For our concerns, the group in question consists of the normal speakers of  $L$ . Condition (ii) is satisfied in spite of the mutual conceptual sufficiency of the two abilities at issue which we have just noted. The ability to think out loud is distinct from the ability to authoritatively say what one thinks. One can have the one without the other. This is true although a normal speaker indeed cannot fail to have both of them (in tandem, as it were). A child who has not yet fully acquired the English language may nevertheless have mastered enough of it to do some thinking out loud (simple things like, *e.g.*, “Daddy is home”). In particular, she may have acquired this ability without having yet mastered the uses of the word “think”.

To see that condition (iii) is satisfied, consider a very crude three-step aetiological picture: Teach a child many things about the use of  $L$ , but not yet the use of “to think that” (or of other linguistic devices which serve similar purposes); then teach or encourage her, if needs be, to soliloquize in your presence (so she learns to think out loud); and in the third step teach her the knack about “I am thinking that ...”, in the sense of this phrase which is pertinent for our topic. This third step amounts basically to the following: Teach her which parts of her speech

productions are such that the result of her prefixing them by “I am thinking that ...” is something true. As soon as the child has learnt this, she is able to say, authoritatively, what she thinks. – What is crucial for our present purposes is not that this simple model comes close to the truth about how the ability to correctly use “I’m thinking that ...” is *actually* acquired. It suffices that this model is coherent and therefore shows a way how, in principle, the ability to authoritatively say what one thinks can be developed by exploiting an antecedent ability to think out loud (and without presupposing an independent ability to ‘find out’, or ‘come to know’ what one is thinking).

As to condition (iv), no equally coherent and persuasive model suggests itself for illustrating how an antecedent authority in *quid*-saying what one thinks may lead to acquiring the ability to think out loud. Hence all four conditions are satisfied. – Moreover, there seems to be no other, independently plausible reason to assume that we could give an explanation ‘the other way round’.

This completes my sketch of an argument for the claim that the ability to think out loud is conceptually explanatory of the ability to authoritatively say what one thinks.

##### 5. Genuine self-knowledge of our thinking unglamorous

Above I had expressed some of my misgivings about the appropriateness of speaking of *knowledge* (instead of *awareness*) to characterize our epistemological position vis-à-vis our conscious thoughts. In order not to create grave misimpressions, let me emphasize here that I do not, of course, deny the plain fact that we can (and often do) have genuine knowledge of our thoughts and of our thinking them. Neither do I deny that this sort of knowledge has several remarkable features not to be found elsewhere.

In the usual course of events, we know what we have just thought and that we have thought it. We know it by simply remembering it. Only in special cases our conscious thinking goes by without leading to knowledge of it – *e.g.*, when we are interrupted or distracted. Most of such knowledge is short-term. Generalizing from my own case I'd say that we tend to forget most of what we have just thought fairly soon. But however short-term it is, as long as such knowledge lingers, it has several characteristics which make it very special. First, there is this striking asymmetry between the one subject who knows what he has just thought and everyone else. Only he can have this piece of knowledge first-hand. Only he can know it 'for sure'. Second, there is this distinctively high reliability. If it seems to him that he has just thought that such-&-such then almost certainly he actually did have that thought. This gives rise to a third noticeable feature: If he tells us what he has just thought and we grant him honesty, then –presuming overall normal conditions– there is almost no toehold for reasonably doubting the truth of what he says. Hence when he tells us what he has just thought, he does so with extraordinary authority.

Clearly then, this is a very special sort of knowledge. And some of the honorific epithets mentioned above can arguably be applied to it (“arrived at via privileged access”, for example, in some sense of this phrase). But equally clearly, it isn't epistemologically glamorous in the sense specified above. Most importantly, it isn't infallible. Furthermore, the extraordinary authority with which one says that he just had this thought isn't supreme authoritativeness. Even the thinker himself to whom it seems that he just had the thought that Harvey is stupid may find reason to doubt that he actually did have this thought.

Well, he could try to enhance his almost perfect certainty (that he gets it exactly right) by employing his ability to think out loud. Here's the thought. Imagine someone who sometimes feels a tiny bit uncertain whether he really remembers the exact thought he's just had. But he is convinced that he remembers some things fully accurately if he hears himself vocalizing them

(telephone numbers, for example). So occasionally, when he wants to be as certain as can be about what exactly he is thinking, he thinks out loud. If he now were to utter “Harvey is stupid”, in full awareness, as only he can be, of the fact that he is honest in doing so (and not, for example, inwardly thinking other things while outwardly reciting some text he knows by heart), he might feel that there is no doubt left for him about what it was that he has just been thinking.

So much is true: He can legitimately dismiss the possibility of his being dishonest in *quod*-saying what he thinks. Nobody else is in a position to do that. For anyone else the possibility of the speaker’s being dishonest is bound to be a relevant alternative. Still, even the knowledge he could acquire in this way would not be immune to any sort of rational (albeit far-fetched) doubt. After all, might he not fear that some nasty device in his surroundings distorts his speech sound, and other things of this kind? So even for him (who cannot doubt his own honesty and now, on hearing himself thinking out loud, trusts his memory unconditionally), there seems to be no way to still all doubts he may reasonably have. Even he would have to *presume* (but arguably cannot know) that nothing interferes.

In brief, genuine self-knowledge of our ‘current’ thoughts is knowledge about the past, however ‘immediate’ the past may be. Therefore, even if supported by additional sensory (audible) evidence, such self-knowledge is not immune to each and every sort of reasonable doubt. It is an extraordinary sort of knowledge. But after all it is fallible, like all other knowledge we have of contingent matters. – The rest is self-awareness.

## 6. Stale glamour

Let’s take stock. I have tried here to find a way of explaining our ability to say authoritatively what we think, without presupposing that we have self-‘knowledge’. The explanation I have

sketched depends on several claims to the effect that certain *ceteris paribus* generalizations are conceptual truths. Such claims are bound to be contentious.

But assuming, ‘for the sake of the argument’, that what I have presented here is on the right track, what would it show about the issue of glamorous self-knowledge? It shows at least one thing, namely that the assumption of such knowledge is not needed. It is not needed in order to account for the most remarkable features characteristic of the special authority with which we can say what we think: infallibility and indubitability. Concerning some of our (first-person) sayings we are infallible, and the truth of what we say cannot be doubted.<sup>19</sup> This, as we have seen, can be conceptually explained in terms of our ability to think out loud. So, there seems to be no reason to assume that we have glamorous self-knowledge, or just knowledge, in order to understand whence we have first person authority concerning what we currently think or why we can speak with supreme authoritativeness on this issue.

The infallibility which is characteristic of saying what we currently think does not require a cognitive achievement. The infallibility at issue is not due to some cognitive condition which enables the speaker to exclude the relevant alternatives. It’s not due to some supremely reliable process of belief-formation concerning the issue *what I am thinking right now*. The speaker requires no reason for believing what he says.<sup>20</sup> Nothing like that (reasons, beliefs, etc.) is needed for being able to say authoritatively what one thinks. All you have to do is to exercise your ability to think out loud, or to say the very thing you would have said, had you acted on this ability.

Concerning the truth of what one says in saying what one thinks, there is, as it were, only one relevant alternative: dishonesty. However, if lack of dishonesty in saying something is *all* that is needed to render unassailably true what a speaker says, then he displays no knowledge in saying it. Honesty is a fine thing, and given favorable conditions it may suffice to guarantee the *truth* of

what is said. About judgments of the “I’m  $\Psi$ ing that  $p$ ” variety –where  $\Psi$  stands for a suitable *verbum dicendi vel declarandi vel intelligendi vel sentiendi vel affectus*– Husserl (1921, II/2, 221) says: “they are indeed true or false, but truth here coincides with honesty [*Wahrhaftigkeit*]”. I am not sure what exactly he means by “coincide” [*zusammenfallen*], but I think he might have been making the same point: With regard to some judgments, honesty is conceptually sufficient for truth. But truth is one thing, and knowledge another. Even if truth may sometimes be assured by honesty alone, knowledge never is.

If some people nevertheless want to speak of self-*knowledge* in this context, it is most probably because they find it hard *not* to speak of a person’s knowing that  $p$ , when the proposition is contingent and the person’s honestly saying that  $p$  conceptually guarantees that  $p$ . For those who are fascinated by the sheer existence of an ability to be (almost) *infallibly* right about *contingent* matters, the uplifting word “knowledge” may seem a meagre minimum of linguistic appropriateness.

The friends of *GSK* will typically ask: How on earth could such a grandiose ability not involve knowledge, especially when acting on it entails getting things right? The short answer suggested by this paper is this, to repeat: If honesty in saying something is *all* it takes for getting certain things right without fail, then knowledge is simply not needed (for getting *these* things right).

Anyway, those who still find it irresistible to speak of knowledge in this connection should also recognize some of the less glamorous aspects of one’s ‘knowledge’ about one’s own current thoughts. Let me, in concluding, mention some of the features I have in mind:

- (a) Such ‘knowledge’ is cheap – in the sense that it is a conceptual free gift which comes with acquiring the status of a normal speaker. (How do you know that you are

thinking that Harvey is stupid? – It would be an answer to say “I have learnt English”.)

- (b) The infallibility of such ‘knowledge’ is extremely ephemeral – the expiry date for this infallibility is the time the next thought occurs. Its glamour lasts exactly as long as the thinking of the thought.
- (c) Items of such ‘knowledge’ are indeed substantial, if only in the sense that they are about contingent facts. But usually they are not very interesting, at least not for other people. Usually, we are not curious about what people just think and wouldn’t give a penny to be told. (Lovers, biographers and psychologists may be exceptions.) What we, occasionally, are interested in is what they think-*and-believe*, or at least think-*and-take-to-be-conversationally-relevant*. Speaking from my own case, most of my conscious sayable thoughts are not worth mentioning. I myself definitely prefer to live in social surroundings in which people are not given to share with me items of their ‘knowledge’ which have this shabby sort of epistemological glamour.

#### 6. Appendant remarks on thought-reports

So far, I have argued that *GSK* is not needed in order to give an account of first person authority. I want now to briefly consider a point which casts doubt on the explanatory effectualness of the *GSK* account itself. It has to do with simultaneity. I have been pretending so far that it makes sense to speak of having and expressing a conscious thought at one and the same ‘exact moment’. This pretence has been made for the sake of argument. I myself am suspicious about the feasibility of this idea. It’s fictitious, I think, nothing but a product of philosophical reveries of super-precision. Clearly it plays no role in how we talk about what we think and when. If in order to let us share his current thoughts, somebody said “I’m thinking right now that so-&-so”, it certainly would be fairly odd, if not just outright silly, to ask him “Well, you said *right now*. What

do you mean? When exactly was it that you had that thought? When you started speaking, just before you started speaking, or as long as you were saying it?”. – This being said, let me continue to ignore such misgivings.

The friends of *GSK* depend on the idea of such simultaneity in their account of the special authority we have in saying what we currently think. Let me briefly sketch the *GSK* picture of what goes on when somebody says “I’m thinking that  $p$ ”. According to it there are three items involved: (a) the thinking that  $p$ , (b) the knowing that one is thinking that  $p$ , and (c) the act of assertively ascribing to oneself the property of thinking that  $p$ . The simultaneity of (a) and (b) is needed for there to be the alleged glamour of the alleged knowledge. For if there were a time lapse, however short, between (a) and (b), memory would become an issue and therefore a gateway for doubts would be opened. The simultaneity of (b) and (c) is needed to secure the special authority the speaker has in making the self-ascriptive utterance. For if there were a time lapse, however short, between (b) and (c), memory would again become an issue, allowing doubts to arise. On this picture, first person authority is crucially tied to the glamour of (b), but this glamour is gone as soon as (a) is past. As soon as the act of thinking that  $p$  is over, the knowledge of one’s thinking that  $p$  has turned, at best, into regular unglamorous knowledge about something in the past. And such regular knowledge would not be explanatory, in the way envisaged by the friends of *GSK*, of the extraordinary authority with which we can say what we think.

We have first person authority in respect of certain types of assertoric present tense self-ascriptions. Given our focus here (on a person’s current thinking) let’s call them *thought reports*. Such a report is an assertive speech act. And it is this which creates a certain difficulty for the *GSK* picture just sketched. In performing an assertive speech act, the speaker means what he says and this involves that he has, or entertains, the thought he expresses by saying what he says. At least in a fully clear case, a speaker who asserts that, *e.g.*, Harvey is stupid has the thought that

Harvey is stupid. (If his utterance were an utterly thoughtless production of speech noise, it would not be an assertion at all. If he thought of something else while uttering “Harvey is stupid”, his performance would at best be a deficient case of asserting that Harvey is stupid.)

So in a non-deficient thought report (“I’m thinking that Harvey is stupid”) the speaker is bound to think that he is thinking that Harvey is stupid. Let us assume –not unreasonably, I think– that at each point in time a normal speaker can consciously entertain at most one sayable thought. Now here’s the difficulty. In the thought report, the speaker ascribes to himself

the thought that Harvey is stupid.

In reporting this, he is to entertain

the thought that he’s thinking that Harvey is stupid.

Hence the thought of which he reports that he is entertaining it is not the thought he is entertaining. And it is not clear that there is any way for him –or anyone else for that matter– to achieve the simultaneity which would be needed, according to the *GSK* picture, for first person authority. At least with respect to *reporting* it ‘in real-time’, the thought I am having right now seems to be an ever-elusive sort of entity.

How does a thinking-out-loud account of first person authority fare with this problem? It may seem that it is inherently hostile to the very idea that there are genuine *reports* about what one is currently thinking (in an artificially strict sense of “currently”). Given the exhibitive nature of thinking out loud, wouldn’t such an account be bound to construe alleged reports of this kind as ‘avowals’, or at least as something which is strictly speaking neither true nor false? No. There are

various options for dealing with ‘real-time’ thought reports. Let me just mention two which can be employed in support of my argument for (3) in section 4 above.

The first says that the report given by assertively uttering

( $\sigma$ ) I’m thinking that Harvey is stupid

is a hybrid speech act in which the subject actually acts on his ability to think out loud. To get the idea, think of a paratactic paraphrase of the sentence uttered:

(P $\sigma$ ) I’m thinking that. Harvey is stupid.

Note that I am not making any claims here about the syntax or semantics of ( $\sigma$ ). Considering this paraphrase at this point is just a preparatory step in sketching the structure of a speech act which can be performed by uttering ( $\sigma$ ). It is meant to make it clearer what I am driving at.

Now here’s the idea. Construe the utterance of the first part of (P $\sigma$ ) as an announcement of an immediately ensuing act of thinking out loud. Construe the utterance of the second part of (P $\sigma$ ) as the performance of the act announced. And consider the report made in uttering ( $\sigma$ ) as the union of these two speech acts. In assertively uttering ( $\sigma$ ) the speaker switches, as it were, from the assertive to the exhibitivite use of language.<sup>21</sup> As to the truth-conditions of such thought reports: What the speaker says, in performing this kind of complex, hybrid speech act is true, given normal conditions, iff he is honest in performing the exhibitivite part of the speech act. As regards the report’s truth condition in abnormal circumstances we keep silent. That is to say, given normal conditions the report is true (as it should be) iff the speaker then thinks that Harvey is stupid.

A second way in which a thinking-out-loud account could be applied to this issue does not view the speaker as actually acting on his ability to think out loud when he utters ( $\sigma$ ). It's enough that he could have done so. The idea is now that what the speaker says in giving the report is true iff he, at the time of the utterance, could just as well have uttered "Harvey is stupid" as a thinking-out-loud. To put it differently: If he had at this moment, instead of uttering ( $\sigma$ ), thought out loud, then "Harvey is stupid" is what he would have uttered. – This is not to say that thought reports really are counterfactuals. We are not concerned here with linguistic analysis or the logical form of thought-report sentences. Rather it is to emphasize that the truth condition of such a self-ascriptive present tense thought report can, for normal circumstances, be specified in a way which is grist to the mill of the thinking-out-loud account. – Both suggestions would need some elaboration.

\*

All these considerations concerned 'real-time' thought reports. Philosophers have a fascination with such reports, but as far as I can tell they are rare in real life. A more realistic way of looking at thought reports would be this. When one is asked at time  $t$  what one is thinking and answers promptly, at  $t'$ , then one standardly tries (and is taken to try) to say what one was thinking at  $t$ . One's answer is usually meant and understood as a report about an event in the immediate past, or in the 'specious present'. As we have seen in section 5, in such a report, the speaker may bring genuine self-knowledge to bear. However, for the reasons pointed out above, such knowledge is not glamorous.<sup>21</sup>

## Footnotes

1 Note that the language is English, actual English, not a language spoken by twin-earthlings. – As a consequence of this explication, if *S* is an entity of which it doesn't make much sense to assume that it could express itself competently in English, then, to the extent in which this doesn't make sense, it is left open what is said in saying that *S* thinks that Harvey is stupid.

2 Here are two examples, more or less randomly selected. According to Tyler Burge, what he calls basic self-knowledge is direct, authoritative, non-empirical and self-verifying (Burge 1988), and in his paper “Our Entitlement to Self-Knowledge” (Burge 1996) he says that *cogito*-like thoughts are direct, self-verifying, infallible and self-evident. Ted Warfield says that knowledge of the contents of our own thoughts is *a priori* and privileged (Warfield 1992, 232).

3 *Cf.*, for example, his *Fourth* and *Fifth Responses* (Descartes 1897–1910, vol. VII, 246, 358 and 359).

4 In a series of recent papers and some unpublished material he was kind enough to make available to me. As to the published papers, *cf.* Dretske 2003a, 2003b and 2004.

5 For philosophical applications of this distinction in the theory of perception and theory of meaning see J. L. Austin (1946, 96 f.) and E. von Savigny (1969, 35 f., 268 f.).

6 Well, maybe not *nobody else*. There are people (“skeptics”) who doubt everything. Or at least philosophers are fond of talking as if there were such people. More interestingly, there are people who conceive of *saying what one thinks* on a translational model. They may object: “Thoughts are mental representations; and in order to say what one thinks, one has to translate one's mental representation in question into a sentence of the respective public language. But translation can go wrong. Hence there is always leeway for doubt. Even if I'm honest in my attempt to tell you what I think, I may have mistranslated my thought.” And there may be even others who feel that they could doubt (c). – Thing is, we don't have to care. For with regard to the dialectical setting of this paper, first person authority is common ground. If those who don't believe in it anyway have misgivings about the proposed *analysans*, why worry? They are –on principle, as it were– not party in our discussion, which is a dispute between two opponents who accept that there is first person authority. After all, my explication is not meant to convince anyone who rejects first person authority. Rather than this it is an attempt to clarify a concept which both the friend of glamorous knowledge and I accept in good faith. – Whatever, you are welcome to read “nobody” as short for “nobody except those who don't believe in first person authority anyway”.

7 I am assuming here that there is a good sense in which, *e.g.*, Birgit who says (at *t*) “I am thinking that Harvey is stupid” *says the same* as Albert when he says “Birgit (at *t*) thinks that Harvey is stupid”. – If you have misgivings about the appropriateness of speaking of same-saying here, you are welcome to replace the phrase, “to say the same” by a more convoluted phrase you take to be apt (*e.g.*, “to say something the truth condition of which is the same as the one of Birgit's saying that she thinks that Harvey is stupid”).

8 A *normal speaker of L* is, roughly, an adult ‘fluent’ in *L*. The crucial thing is that he has a high degree of mastery of *L*: speakers of *L* who use *L* as their common language would consider him as of equal linguistic competence. – So, being a normal speaker of *L* turns on what kind of linguistic *ability* a person has. (It is not required that he is an actual member of an *L*-speaking community. Nor is it required that he is a normal person; being linguistically competent to the relevant degree is compatible with having foibles, quirks and abnormalities.)

9 The *cp*-clause must of course not be understood as an *ad hoc* device for immunizing the generalization against refutation (“except when it’s not so”). For more on *cp*-necessities see Kemmerling (2006, 122 f. and 133 ff.).

10 It should be noted that what is thought out loud must be a thought. This is to say that what is uttered, in a case of thinking out loud, has to be something which is syntactically and semantically appropriate to complete the lacuna of “the thought that ...”. Thinking out loud is indeed a kind of monological talking which is not audience-directed (or, if you wish, a kind of soliloquizing, even if performed in the presence of an audience). But not each and every kind of speech noise which is produced in this manner constitutes a thinking out loud, in the sense in which I am using this term here. The propositional wheat has to be separated from the chaff of the utterance. – This restriction is unpleasantly strong, in at least one sense: many meaningful parts of what is said in this monological manner might have to be excluded. Questions, exclamations, imperatives, etc. should also be considered as linguistic forms in which some of our sayable thoughts can be revealed exactly as we entertain them. A full-fledged theory of thinking out loud should cover such thoughts too. Lacking such a theory, I grudgingly exclude them. – I hasten to add that none of this affects my case against glamorous self-knowledge.

11 In the next section I shall return to the issue of the conceptual connection between thinking out loud and honesty.

12 Notice the word “pure”. – It might be worth mentioning in passing that *impure* acts of thinking out loud are much more common. Performing regular speech acts of various sorts, in what Ryle calls *normal unstudied talk*, can be a way of thinking out loud. (See Ryle 1949, 181 ff.) But it is characteristic of many such ‘impure’ acts that we are aware of the fact that we commit ourselves, in making the utterance, to holding true the thought which we put forward, and sometimes even to the thought’s truth. – In this paper, I shall use “thinking out loud” as a short way of referring to a pure act of thinking out loud in the presence of an audience (where it is usually common knowledge between speaker and audience that such an act is being performed).

13 “It is the defining characteristic [of a declaration] that the successful performance of one of its members *brings about* the correspondence between the propositional content and reality, successful performance guarantees that the propositional content corresponds to the world ... Declarations *bring about some alteration* in the status or condition of the referred to object or objects solely in virtue of the fact that the declaration has been successfully performed.” (Searle 1979, 16 f.)

14 Concededly, “thinking out *loud*” may not be appropriate for describing anything which is done in painting.

15 Fred Dretske straightened me out on this in a lecture he gave in Heidelberg 2008.

16 To be fully accurate, I should say “*cp*-conceptually sufficient”, but I trust no-one will begrudge my not always marking this distinction. The reader should know by now that I consider almost any interesting sort of conceptual relationship as one in which overall normality is essentially involved.

17 But think of a typical learner’s questions. A novice in chess may ask “Why has Black won this game?” Answers like “Because Black checkmated White” or “Because Black challenged the white king in such a way that he could not avoid being captured” may give him *exactly* the sort of explanation which was hoped for. More ‘informative’ or ‘substantial’ answers (like “Because from

the 12<sup>th</sup> move on, White was not able to prevent the promotion of Black's free pawn") may not yet be helpful.

18 Conceptual explanation appears to be a fairly neglected topic. But see Benjamin Schnieder (Schnieder 2006a, esp. 402–411; 2006b, esp. 31–39) for some interesting thoughts.

19 I am not so sure about the applicability of some of the other labels to be found in the literature. But "*a priori*" definitely is a horrid misnomer. The alleged knowledge is neither about something which is true *a priori*, nor can it be had independently of experience. For, first, what one says when one says what one currently thinks is not an *a priori* truth. Second, one couldn't come to know what one then allegedly knows without the experience of having the thought in question.

20 In a way, as we have noticed in section 4, the fact that he says what he says, and hears himself saying it, may give him a reason to believe that what he says is what he is thinking.

21 Beware of taking "switch" too literally. This word is definitely not meant to suggest any psychological reality. Compare what one does in the course of saying something like "Harvey gets angry with Birgit whenever she tickles him": he variously switches from using words referentially to using (other) words predicatively. The switch, in uttering ( $\sigma$ ), from asserting to exhibiting would have about the same sort of reality as the switches from referring to predicating in uttering a sentence like the one just mentioned.

22 Thanks to Steven Davis, Fred Dretske, Wolfgang Freitag, Sanford Goldberg, Stefanie Grüne, Klaus Host, Adam Leite, Felix Mühlhölzer, Friederike Schmitz and Marcus Willaschek for their help in finally bringing about the final version. Special thanks to Mark Helme.

## References

- Austin, J. L., 1946: Other Minds, *Proceedings of the Aristotelian Society* 20, 148–87, quoted after the reprint in *Philosophical Papers*, ed. J. O. Urmson and G. J. Warnock, Oxford: Oxford University Press<sup>2</sup>1970, 76–116
- Burge, T., 1988: Individualism and Self-Knowledge, *Journal of Philosophy* 85, 649–663
- Burge, T., 1996: Our Entitlement to Self-Knowledge, *Proceedings of the Aristotelian Society* 96, 91–116
- Descartes, René, 1897–1910: *Œuvres*, ed. C. Adam and P. Tannery, Paris: Cerf
- Dretske, F., 2003a: How do you know you are not a zombie, in: *Privileged Access and First Person Authority*, ed. B. Gertler, Burlington: Ashgate, 1–13
- Dretske, F., 2003b: Externalism and self knowledge, in: *New Essays on Semantic Externalism and Self-Knowledge*, ed. S. Nuccetelli, Cambridge, Mass.: MIT Press, 131–142
- Dretske, F., 2004: Knowing what you think *vs* knowing that you think it, in: *The Externalist Challenge*, vol. 2, ed. Richard Schantz, Berlin/New York: de Gruyter, 389–399
- Husserl, E., 1921: *Logische Untersuchungen*, 2 volumes, 2<sup>nd</sup> ed., Halle: Niemeyer
- Kant, Immanuel, 1787: *Kritik der reinen Vernunft*, 2<sup>nd</sup> ed., Riga: J. F. Hartknoch
- Kant, Immanuel, 1800: *Anthropologie in pragmatischer Hinsicht*, 2<sup>nd</sup> ed., Königsberg: F. Nicolovius
- Kemmerling, A., 2006: Kripke's Principle of Disquotation and the Epistemology of Belief Ascription, *Facta Philosophica* 8, 119–143
- Ryle, G. 1949: *The Concept of Mind*, London/New York: Hutchinson's University Library
- von Savigny, E., (1969): *Die Philosophie der normalen Sprache*, Frankfurt a.M.: Suhrkamp
- Schnieder, B. (2006a): A Certain Kind of Trinity: Dependence, Substance, Explanation, *Philosophical Studies* 129, 393–419
- Schnieder, B. (2006b): Truth-making without Truth-makers, *Synthese* 152, 21–46
- Searle, J. R. (1979): *Expression and Meaning*, Cambridge: Cambridge UP
- Warfield, T. (1992): Privileged Self-Knowledge and Externalism Are Compatible, *Analysis* 52, 232–237