TABLE OF CONTENTS

Special Topic I
WITTGENSTEIN

Guest Editors
Kai BÜTTNER, Florian DEMONT, David DOLBY, Anne-Katrin SCHLEGEL

Special Topic II
QUINE

Guest Editor
Dirk GREIMANN

Special Topic

WITTGENSTEIN

Grazer Philosophische Studien
89 (2014), 3–10.

INTRODUCTION

Kai BÜTTNER, Florian DEMONT, David DOLBY,
Anne-Katrin SCHLEGEL

University of Zurich

Mathematics occupied a central place in Wittgenstein's work. He was led to philosophy by an interest in disputes about the foundations of mathematics. His reflections resulted in the *Tractatus*' short but insightful critique of the logicism of Frege and Russell. Moreover, it is reported that he decided to return to philosophy in 1929 after having attended a talk given by the intuitionist mathematician Brouwer. In the 1930s and early 1940s Wittgenstein wrote extensively on the philosophy of mathematics, covering a large variety of themes, including the notions of number and infinity, the method of proof by induction, the role of contradictions and consistency proofs in mathematics, the application of mathematics, and the nature of mathematical proof and mathematical necessity. In 1944 he stated that his most significant contribution had been in the philosophy of mathematics (Monk 1990, 466).

However, the initial reception of Wittgenstein's philosophy of mathematics was predominantly negative (Dummett 1959, Kreisel 1958, Bernays 1959). Many of his remarks about mathematics were dismissed as either wrong or irrelevant; and some were alleged to contain technical errors. To a certain extent, this reaction can be explained by the radical nature of Wittgenstein's ideas, which often challenge traditional assumptions about mathematics. Another cause has been misunderstanding: the interpretation of Wittgenstein's writings has proven difficult due to his unconventional style of writing, the fact that the material from his *Nachlass* is partly unfinished and was not intended for publication, and the delayed publication of many relevant manuscripts from his middle period. Fortunately, the exegesis of Wittgenstein's writings on mathematics has progressed considerably in recent years, with the appearance of such excellent monographs as Shanker 1987, Frascolla 1994, Marion 1998 and Mühlhölzer 2010. These books clarify Wittgenstein's claims and show

how the accusations of technical incompetence were often the result of misinterpretation. This has led to a greater appreciation of Wittgenstein's unorthodox views, which promise fresh perspectives on debates that had appeared to have reached deadlock. The essays in this volume are recent contributions to the newly invigorated debate about Wittgenstein's philosophy of mathematics by leading scholars and specialists.

According to the later Wittgenstein, many philosophical misconceptions and mythologies can be traced back to the assumption that any declarative sentence describes a corresponding portion of reality. Thus, he famously argues that self-ascriptions of mental predicates have an expressive function and do not describe an inner world of private experiences. And, similarly, mathematical propositions have a normative rather than a descriptive function. Instead of describing a realm of abstract objects, they are rules for the transformation of empirical descriptions. In his paper Pasquale Frascolla discusses a potential difficulty for Wittgenstein's conception, which arises from the intuitive assumption that for a sentence to have a descriptive function is equivalent with its being truth-apt. According to Frascolla, Wittgenstein did not mean to deny this principle in order to make room for sentences which are both non-descriptive and truth-apt. Instead he considered neither mathematical propositions nor self-ascriptions of mental predicates to be truth-apt. Mathematical propositions, in particular, correspond to reality at best in the way in which rules correspond to social practices. And the proof of a mathematical proposition determines the latter's sense rather than establishing its truth.

In his contribution Severin Schroeder investigates how we should understand Wittgenstein's claim that mathematical propositions are rules of grammar for non-mathematical language. In particular he explores how to resolve the tension between Wittgenstein's claim that mathematical propositions are rules and his emphasis on the practical usefulness of mathematics. As Wittgenstein himself asks: "How can the mere transformation of an expression be of practical consequence?" (RFM 357) According to Schroeder, Wittgenstein's answer is that mathematical propositions are not merely stipulative definitional rules: they forge connections between independently comprehensible concepts. A proposition expressing an empirical correlation may be fixed as a rule. It is in this way that mathematical propositions are "dependent on experience but made independent of it" (LFM 55). A problem for Wittgenstein's view nevertheless remains: what violates a grammatical rule is nonsense and yet non-mathematical statements that are not in conformity with mathematical propositions would not be nonsensical but rather false (except in the most simple cases). Schroeder's solution is to suggest that mathematical propositions do not determine what makes sense in natural language: they are norms for the activity and discourse in which we develop and apply a system for calculating quantities (rather than simply counting or measuring them).

Felix Mühlhölzer addresses a problem which has occupied philosophers as well as mathematicians. It comes from elementary results in model theory to the effect that many formalized theories, e.g. first-order Peano Arithmetic, have non-standard models. On the one hand, if one gives a precise formulation of what arithmetic is about, as is done by the model-theoretic notion of interpretation, then one cannot catch it uniquely: there is a multitude of non-intended interpretations that one cannot dismiss. If, on the other hand, one specifies the standard model (i.e. the intended interpretation) in the meta-language in which model-theory is (often informally) expressed, then what arithmetic is about is not made precise and transparent. Mühlhölzer calls this the *aboutness dilemma*. He argues that this dilemma can be seen as rooted in a blurring of the categorial difference between the notions of 'interpretation' and 'reference'. The interpretations mentioned in the first horn are simply mathematical functions that do not involve any use of the so-called 'signs' that are interpreted. These signs are 'dead', 'petrified', and one should not think that petrification is simply idealization: idealization abstracts from non-essential aspects; petrification, by contrast, from the essential ones, namely aspects of use. The second horn of the aboutness dilemma concerns 'reference', a notion which Wittgenstein claims is essentially tied to the use of signs (see PI § 10). This view adopts what one might call the *full use thesis*, according to which there is nothing more to 'reference' than what can be seen in the use of our terms, and rejects the *partial use thesis*, usually adopted in the literature, which claims only that the use of our terms contributes to the necessary link between ourselves and the objects referred to. From the Wittgensteinian point of view, then, one can accept the first horn of the aboutness dilemma, as asserting an uncontroversial mathematical result, and correct the second horn by presenting a perspective on 'reference' that is sufficiently clear (even though imprecise).

One of the constant targets of Wittgenstein's criticism is Frege and Russell's logicism. In their paper "Wittgenstein on equinumerosity and surveyability" Marion and Okada try to connect two of Wittgenstein's objections, which they call the 'modality argument' and the 'surveyability argument'. The former is directed against the logicist's definition of equinumerosity according to which two concepts F and G are equinumerous if and only if there is a one-one relation between the Fs and Gs. Wittgenstein objects to this definition that the possibility of a one-one correlation between the Fs and Gs is a consequence of, rather than a condition for, the equinumerosity of F and G. For it is only when the numbers of the F and G are comparatively small that one can determine whether F and G are equinumerous without counting the Fs and Gs and hence without identifying their respective numbers. In most cases, however, the equinumerosity of F and G needs to be derived from the identity of their numbers. The surveyability argument, on the other hand, concerns proofs in Russell's *Principia Mathematica*. According to Wittgenstein, only a very limited portion of arithmetic can actually be proven within this system. Due to its complicated notation, proof-constructions within this system quickly become unsurveyable and thereby loose their cogency. Marion and Okada suggest that both arguments draw on the same type of consideration: those concerning the limits of visual thinking. For the crucial premise of the surveyability argument is that only perspicuous sign-patterns can count as proofs. And the modality argument presupposes that the possibility of determining the equinumerosity of F and G without counting depends on whether or not the Fs and Gs can be appropriately visualized, as, for example, by being arranged in a surveyable pattern.

Esther Ramharter observes that although Wittgenstein frequently discusses *equations* (e.g. $x^2 + 1 = 0$, $\bigwedge_x x(x-1) = x^2 - x$) and *formulae in a narrow sense* (e.g. sin $2x$, $r^2\pi$), he rarely discusses the two together, let alone discusses the latter as parts of equations. She raises two issues for which this is of consequence: First, in the *Philosophical Remarks* §§ 166–167 Wittgenstein contradicts himself by claiming: (i) that only a proof can give a mathematical proposition sense; (ii) that induction is not a proof; and (iii) that algebraic propositions gain their sense through induction. Ramharter argues that this contradiction arises because, according to Wittgenstein, the meaning of a formula in the narrow sense and the meaning of a proposition are given in distinctly different ways. In the case of algebraic equations, which are mathematical propositions but are composed of formulae in a narrow sense, those two meanings clash. She concludes that the need to broaden the concept of the meaning of a formula, as Wittgenstein later did in the *Remarks on the Foundation of Mathematics*, had already been apparent in the early 1930s. Secondly, Wittgenstein's insistence that the sense of a mathematical proposition is determined by its proof seems to leave no room for mathematical problems. In his later work, however, Wittgenstein was explicit that there are mathematical problems. But how they are to be conceived on his account is unclear. Ramharter argues that an explanation compatible with Wittgenstein's remarks—and even presupposed by some of them—is available once the correct account of formulae in a narrow sense as parts of equations is in place.

The extent of agreement between Wittgenstein's philosophy of mathematics and the intuitionism of Brouwer is contentious. In his contribution Ian Rumfitt argues that Wittgenstein accepted one of Brouwer's main claims: some of the higher mathematics of their day rested upon an illegitimate projection into the infinite of methods that properly apply only within finite domains. He goes on to compare and assess the different treatments the two philosophers give of problematic cases involving infinity. According to Brouwer, an exception to the law of excluded middle is provided by the proposition "There are three consecutive sevens in the expansion of pi". This proposition is not guaranteed to have a truth-value because an infinite series can have no properties over and above those entailed by the generating principles themselves, and it is not guaranteed that we can derive from these principles a proof that the sequence does or does not occur in the expansion. Therefore, we cannot assert "Either there are or there are not three consecutive sevens in the expansion". Wittgenstein, by contrast, denies that such cases are exceptions to the law of excluded middle: for him the applicability of the law is a criterion of being a proposition. He argues instead that such propositions lack sense. Rumfitt identifies two considerations Wittgenstein gives in favour of this view and then outlines possible responses for the intuitionist.

Earlier versions of these papers were presented in August 2012 at the conference "Perspectives on Wittgenstein's Philosophy of Mathematics", organised by the Chair of Theoretical Philosophy at the University of Zurich (Hans-Johann Glock) with the financial support of the Swiss National Science Foundation and the Marie Gretler Foundation. We would like to thank Adrian Frey and Susanne Huber for their excellent work in organizing the conference with us; and to Göran Sundholm and Pieranna Garavaso who presented their research at the conference and made invaluable contributions to discussion.

References

Bernays, Paul 1959: "Betrachtungen zu Ludwig Wittgensteins Bemerkungen über die Grundlagen der Mathematik". *Ratio* 2, 1–22.

Dummett, Michael 1959: "Wittgenstein's Philosophy of Mathematics". *Philosophical Review* 68, 324–48.

Frascolla, Pasquale 1994: *Wittgenstein's Philosophy of Mathematics*. London: Routledge.

Kreisel, Georg 1958: "Wittgenstein's Remarks on the Foundations of Mathematics". *British Journal for the Philosophy of Science* 9 (34), 135–158.

Marion, Mathieu 1998: *Wittgenstein, Finitism, and the Foundations of Mathematics.* Oxford: Oxford University Press.

Monk, Ray 1990: *Ludwig Wittgenstein: The Duty of Genius*. London: Vintage.

Mühlhölzer, Felix 2010: *Braucht die Mathematik eine Grundlegung?: Ein Kommentar des Teils III von Wittgensteins Bemerkungen über die Grundlagen der Mathematik*. Frankfurt: Klostermann.

Shanker, Stuart 1987: *Wittgenstein and the Turning Point in the Philosophy of Mathematics.* New York: SUNY Press.

Abbreviations for Works by Wittgenstein

AWL	*Wittgenstein's Lectures, Cambridge, 1932–35*. A. Ambrose (ed.). Oxford: Blackwell, 1979.
BB	*The Blue and Brown Books*, 2nd edition. Oxford: Blackwell, 1969.
BT	*The Big Typescript: TS 213*. German-English scholars' edition. C. Grant Luckhardt & Maximilian A. E. Aue (tr.); C. Grant Luckhardt & Maximilian A. E. Aue (eds.). Oxford: Blackwell, 2005.
LC	*Lectures and Conversations on Aesthetics, Psychology and Religious Belief.* Cyril Barrett (ed.). Oxford: Blackwell, 1966.
LFM	*Wittgenstein's Lectures on the Foundations of Mathematics, Cambridge 1939*, from the notes of R. G. Bosanquet, N. Malcolm, R. Rhees and Y. Smithies. Cora Diamond (ed.). Ithaca, New York: Cornell University Press, 1975.
LWL	*Wittgenstein's Lectures*, Cambridge*, 1930–32*, from the notes of J. King and D. Lee. D. Lee (ed.). Oxford: Blackwell, 1980.
MS	Manuscripts from Wittgenstein's Nachlass, published as *Wittgenstein's Nachlass: The Bergen Electronic Edition*. Oxford: Oxford University Press, 2000.

NB *Notebooks, 1914–16*. G.E.M. Anscombe (tr.); G.H. von Wright & G.E.M. Anscombe (eds.). Oxford: Blackwell, 1961.

OC *On Certainty*. G.E.M. Anscombe & G.H. von Wright (eds.). Oxford: Blackwell, 1969.

PG *Philosophical Grammar*. Anthony Kenny (tr.); Rush Rhees (ed.). Oxford: Blackwell, 1974.

PI *Philosophical Investigations*. 3rd edition. G.E.M. Anscombe (tr.); G.E.M. Anscombe & R. Rhees (eds.). Oxford: Blackwell, 1953. Revised 4th edition. G.E.M. Anscombe, P.M.S Hacker & Joachim Schulte (tr.); P.M.S. Hacker & Joachim Schulte (eds.) Oxford: Blackwell, 2009.

PLP *The Principles of Linguistic Philosophy*. By F. Waismann. R. Harré (ed.). London: Macmillan, 1965.

PPO *Ludwig Wittgenstein: Public and Private Occasions*. J. Klagge & A. Nordmann (eds.). Lanham MD: Rowman & Littlefield, 2003.

PR *Philosophical Remarks*. R. Hargreaves & R. White (tr.); R. Rhees (ed.). Oxford: Blackwell, 1975.

PUKG *Philosophische Untersuchungen, kritisch-genetische Edition*. Joachim Schulte (ed.). Frankfurt a. M.: Suhrkamp, 2001.

RFM *Remarks on the Foundations of Mathematics*, 3rd edition. G.E.M. Anscombe (tr.); G.H. von Wright, R. Rhees & G.E.M. Anscombe (eds.). Oxford: Blackwell, 1978.

RPP I *Remarks on the Philosophy of Psychology*, Volume I, G.E.M. Anscombe (tr.); G.E.M. Anscombe & G.H. von Wright (eds.). Oxford: Blackwell, 1980.

TLP *Tractatus Logico-Philosophicus*. D.F. Pears & B.F. McGuinness (tr.). London: Routledge and Kegan Paul, 1961.

WVC *Wittgenstein and the Vienna Circle,* from the notes of F. Waismann. Brian McGuinness (ed.); Joachim Schulte & Brian McGuinness (tr.). Oxford: Basil Blackwell, 2003.

Z *Zettel.* G.E.M. Anscombe (tr.); G.E.M. Anscombe & G.H. von Wright (eds.). Oxford: Blackwell, 1967.

List of Contributors

Pasquale FRASCOLLA is Professor of Philosophy at the University of Basilicata. He is the author of *Wittgenstein's Philosophy of Mathematics* (1994) and *Understanding Wittgenstein's Tractatus* (2011), as well as of numerous papers on the philosophy of mathematics, the *Tractatus*, and inductive rationality.

Mathieu MARION is Professor for Philosophy and Logic at the University of Québec in Montréal. He is the author of *Wittgenstein, Finitism and the Foundations of Mathematics* (1998) and has made seminal contributions to the investigation of Wittgenstein's philosophy of mathematics during his so-called middle period.

Felix MÜHLHÖLZER is Professor of Philosophy at the Georg-August University in Göttingen. He has published several papers in the philosophy of mathematics and has written a commentary on Part III of Wittgenstein's 'Remarks on the Foundations of Mathematics' *Braucht die Mathematik eine Grundlegung?* (2010).

Mitsuhiro OKADA is Professor of Philosophy at Keio University and project leader of the Logic and Language Lab at the Centre for Integrated Research on the Mind (CIRM) of Keio University. He has written several papers in the areas of philosophical logic and proof theory, theoretical linguistics and the philosophy of mathematics.

Esther RAMHARTER is Associated Professor of Philosophy at the University of Vienna. She is author of *Die Härte des logischen Muß. Wittgensteins "Bemerkungen über die Grundlagen der Mathematik"* (2006) and editor of *Prosa oder Beweis? Wittgensteins "berüchtigte" Bemerkungen zu Gödel* (2008).

Ian RUMFITT is Professor of Philosophy at Birkbeck College, University of London. He has published widely on the philosophy of logic, intuitionism and the history of analytic philosophy.

Severin SCHROEDER is Reader in Philosophy at the University of Reading. He is the author of *Wittgenstein: The Way Out of the Fly-Bottle* (2006) and *Wittgenstein Lesen* (2009). He has published several articles on rule-following and on Wittgenstein's account of analytic truth.

Grazer Philosophische Studien
89 (2014), 11–21.

REALISM, ANTI-REALISM, QUIETISM: WITTGENSTEIN'S STANCE

Pasquale FRASCOLLA
University of Basilicata, Italy

Summary

I shall defend the view that, according to Wittgenstein, mathematical propositions are not truth-apt. With respect to empirical propositions, Wittgenstein takes a mild anti-realist stance based on the identification of truth with warranted assertibility. As regards mathematical propositions, however, his position amounts to a form of a radical anti-realism, understood as the denial of the possession of assertoric content. Provability is the only legitimate *Ersatz* for truth in the case of mathematical sentences. But Wittgenstein claims that a proof is a means not to establish the truth of a mathematical proposition but to determine its sense.

In a famous passage of the *Philosophical Investigations*, Wittgenstein clearly states that the recognition of the diversity of functions that language performs is an indispensable condition for freeing oneself from the philosophical paradoxes that arise from overlooking the differences of use in its various regions: "The paradox disappears only if we make a radical break with the idea that language always functions in one way, always serves the same purpose: to convey thoughts—which may be about houses, pains, good and evil, or anything else you please" (PI § 304) (and one could add: about numbers, even though Wittgenstein does not mention them explicitly). To convey thoughts, i.e. to convey doxastic and epistemic contents, is just one among the many things that can be done by using language, one among the many language games that can be played, and is what is usually done by using declarative sentences with assertoric force. The joint command of the "designation-object", or "name-bearer", model over the representation of the grammar of words (singular terms, common nouns etc.), and of the descriptivist model over the representation of the grammar of declarative sentences (a sentence of that kind describes a state of affairs which, if it obtains, makes the sentence true), leads to deep misunderstandings of the effective conditions of linguistic usage and generates paradoxes, mysteries, philosophical mythologies. Since a claim to the obtaining of the state of affairs that a declarative sentence describes is made by asserting it, and thus a claim to the truth of the sentence, transmission of doxastic and epistemic contents, practice of assertion and descriptivist, or truth-conditional model in the representation of grammar, appear to be bound together in one lump.

In that long tradition of Wittgensteinian studies that Monk has called "the right wing", Wittgenstein's warning not to overlook the differences in function that are hidden under the syntactic uniformity of surface, has been taken strictly[1]. The interpretation I have in mind runs along the following lines: there are regions of discourse whose sentences, in spite of their misleading syntactic appearance, do not play the role of describing a corresponding portion of reality; as a consequence, those sentences do not convey any doxastic or epistemic content and do not belong to the field of what is assertible: in short, they are not truth-apt. This drastic conclusion holds of several classes of sentences (the *critical classes*, as I shall be calling them from now onwards). First of all, it holds of mathematical sentences, which Wittgenstein often invites us to imagine as framed not in the indicative mood, with the usual assertoric force attached to them and as conveying contents of propositional attitudes like belief or knowledge, but in the imperative mood, as conveying orders. On nearly every occasion in Wittgenstein's writings, mathematical sentences are compared with the expressions of norms, of grammar rules, and are construed as parts of the apparatus of language, tools for constructing descriptions, not descriptions themselves. This is the picture that he more frequently opposes to the Platonist view that mathematical sentences describe an ethereal reality of ideal objects, causally inert but, at the same time, capable of providing the cognitive claims of mathematics with a firm foundation (famously, "the alchemy of mathematics").

Self-ascriptions of qualitative states or, more generally, of mental states, undergo an analogous treatment when Wittgenstein construes them as avowals and not as statements. Here the distinction between third-person ascriptions, used with assertoric illocutionary force and endowed with cognitive import, and first-person ascriptions, playing an expressive role, aims at releasing the representation of their grammar from the assump-

1. See Monk 2007.

tion of a scenario of inner objects and events, and at freeing us from the venerable philosophical problems rising from the sharp contrast between the epistemic authority of the first person and the fallibility of the third person. As Simon Blackburn has often stressed, similar conclusions hold, according to Wittgenstein, of ethical and aesthetical judgements, of expressions of religious belief, of the hinge-propositions of common sense, with specific traits in each case[2].

Since the late Eighties, however, the whole picture began changing as a result of the convergent interventions of authors as distant as can be in their overall approach, if we think of Sabina Lovibond, the new Wittgenstein duo Cora Diamond and James Conant plus Hilary Putnam, Crispin Wright and Paul Horwich. Those among them who can be classified as Wittgenstein scholars, belong to the "left wing", according to Monk's partition. A thorough reflection on Wittgenstein's meta-philosophy, on the goals and methods of grammatical enquiry with respect to the philosophical tradition and its problems, and on the purported quietist implications of that enquiry, is, at least, a central theme of the investigations of all those authors. Here I will seek to throw light on the impact that their work has had on our problem, i.e. on how Wittgenstein's appeal to take into account the differences between the various regions of discourse should be understood, especially as concerns his reflections on mathematics and the realism/anti-realism controversy. A first contention is that a sort of un-pretentious realism (Lovibond 1983), or a common-sense realism (Putnam 1996, 2001, 2008), can find inspiration and support in Wittgenstein's writings. This variety of realism should not be confused with philosophical realism, to which anti-realism, as an alternative philosophical thesis, is opposed: Wittgenstein would not take sides in favour of one of the contenders, but would show how the conflict itself evaporates once the mistaken grammatical assumptions shared by the two parties are recognized and dismissed. A second, more refined view stems from Wright's minimalism, according to which the controversy between realism and anti-realism is to be placed in a new common theoretical framework, accepted by both the contenders. That framework is characterized by the presence of a light notion of truth, on which the notion of truth-aptness of assertoric contents is based, where assertoric contents, in turn, are defined in terms of syntax and discipline of justification (Wright 1992, 2003).

2. See Blackburn 1990, 1998, 2010.

As a matter of fact, the positions within this field are often diverging on central issues. Lovibond, for instance, has a view that, quite surprisingly, leads her to attribute to Wittgenstein a conception of language that denies differences altogether. Wittgenstein's would be a homogeneous or seamless conception of language, based on the idea that sentential syntactic form does not deceive us: if a linguistic expression has the form of a declarative sentence in the indicative mood, we have to treat it as descriptive or fact-stating, because the descriptive function pervades all regions of discourse irrespective of differences in content. The generalized application of the descriptivist model goes hand in hand with the attempt to purify it of all its realistic implications, through the adoption of a redundancy theory of truth (as can be found in the *Philosophical Investigations*, § 136, and in many passages of the *Remarks on the Foundations of Mathematics*), and the neutralization of any purported metaphysical import of the notion of aboutness. From that point of view, the status of sentences that are about specific regions of reality and that are used to make assertions on them can be conferred both on ethical and mathematical sentences.

It seems to me that Lovibond's is clearly untenable as an interpretation of Wittgenstein's stance. Instead of going into the details of her position, I will focus on some traits shared by the views of the authors I am considering, in order to clarify their overall strategy and to compare it with a seemingly similar strategy that, in point of fact, can be found in Wittgenstein's writings, especially in those passages where a quietist stance, at least apparently, is espoused. A quick comparison with Wright's minimalism may be expedient in this connection. Wright's starting point in many respects recalls Lovibond's: to have an assertoric content is not a deep feature of some declarative sentences that can be mimicked by the syntactic form of other sentences, which would possess it only apparently. For both Wright and Dummett, a similar point can be made as regards reference to objects: once a linguistic expression fulfils the logical and syntactic requirements for being a singular term, no further question can be raised about whether or not it succeeds in referring to an object, on condition that it occurs in statements of predication and in identity statements recognized as true by ordinary criteria (according to Wright, this conception of reference would be at the basis of Frege's conception of numbers as objects). In a perfectly analogous way, there is no strong notion of "genuine" assertoric content such that its application would lead to the conclusion that the sentences of a certain class cannot be credited with the capability of conveying an assertoric content, in spite of the fact that they are subject to certain minimal constraints of syntax (being constituents of truth-functions and occurring as that-clauses in ascriptions of propositional attitudes, among them), and belong to a discourse that is disciplined by standards of warranted assertibility. Once a properly weakened notion of a sentence endowed with assertoric content has been adopted, truth-aptness can be conferred on all the sentences possessing that content, provided that the truth predicate is characterized, in turn, by means of metaphysically non-committal principles and rules (the Disquotational Schema, for instance), as minimalism actually does.

Wright has suggested that some of Wittgenstein's statements could be understood in a way that makes them consistent with minimalism: among them, his (Wittgenstein's) appeal to the Disquotational Schema within a deflationary conception of truth and the recurring characterization of propositions as elements of the truth-functional calculus. As Wright himself acknowledges, Wittgenstein, given his non-constructive view of the task of philosophy, was not interested in developing an overall theoretical framework in which the traditional conflict between realism and anti-realism could be framed in new terms. In my opinion, Wright's suggestion is definitely more convincing than Lovibond's approach: the attribution of an assertoric content to a sentence on minimalistic grounds does not derive, as it does in Lovibond's interpretation of Wittgenstein, from the easily questionable attribution of a descriptive function, albeit metaphysically watered-down, to the sentences of all regions of discourse, mathematics included.

The reading of Wittgenstein's philosophy of mathematics proposed by Diamond, Putnam and Conant makes a further step in the direction of an interpretation of the theme of the differences that entails no radical opposition between the sphere of assertibility and truth-aptness, on the one hand, and that of the expression of rules, on the other. The way they try to achieve their goal, however, is very different from both that of Lovibond and that of Wright. According to Diamond and Conant, Wittgenstein was actually engaged in stressing the variety of functions that declarative sentences perform in the various regions of discourse (a descriptive function, an expressive function, a normative function, and so on), quite independently of their superficial syntactic uniformity or, better, contrary to what that uniformity misleadingly suggests: both Lovibond and Wright would lend too great a weight to syntax and too little a weight to use, which is the true final court of appeal for questions of meaning. The crucial tenet in Conant's reading is that acknowledging a non-descriptive role of the sentences of a certain class does not amount to ruling them out of the sphere of assertibility and hence of truth-aptness. Conant is quite explicit on this point: speaking of self-ascriptions of mental states, he declares that the goal of philosophical enquiry is that of "obtaining a perspicuous overview of the interplay between the various functions of avowals (among which are its expressive and assertoric functions)" (Conant 1996, 207). Talking about mathematical sentences, he maintains that in order to unmask the nature of pseudo-problems of traditional philosophical riddles, the acknowledgement of the interplay of the multiple functions mathematical sentences perform is called for: "Wittgenstein's treatment of mathematics can in this respect be seen in its general approach to parallel his treatment of avowals" (Conant 1996, 221). Conant's strategy is clear: room has to be made for the notion of a non-descriptive statement, in order to account both for the use of mathematical sentences in the traffic of argument and inference, in the practice of giving and asking for reasons—this is done by classifying them as statements—and for the normative role they usually play—this is done by qualifying them as non-descriptive.

To be honest, it seems to me that Conant's is a mere verbal trick, good to have it both ways. I will try to briefly outline a defence of the right-wing interpretation of the theme of differences, focussing on some aspects of Wittgenstein's anti-realism in the philosophy of mathematics. The decisive question is: what is Wittgenstein's purpose when he invokes the Disquotational Schema and espouses the redundancy conception of truth, and when he gives his weak characterization of the notion of a proposition as an element of the truth-functional calculus? In my opinion, these are not to be construed as the premises of an argument leading to the inclusion of the sentences of the critical classes (mathematics, folk-psychology, ethics, aesthetics etc.) in the realm of assertibility and truth-aptness. On the contrary, they are the means for liberating the notion of truth from any realistic mortgage and for depriving it of any additional content beyond that which the notion of warranted assertibility, or justifiability according to socially accepted standards, has of its own. The anti-realistic conception of truth, however, applies only within the doxastic and epistemic sphere, that of empirical sentences, of genuine descriptive statements, whereas the sentences of the critical classes, and mathematical sentences in particular, remain outside that field of application.

The textual evidence in favour of the above conjecture is well known, which allows me to confine myself to recalling just a couple of things. First of all, there is what Wittgenstein says of truth in *On Certainty* (§§ 199–206), where the identification of truth and warranted assertibility is overtly maintained in relation to empirical propositions, with a view to weakening the realistic implications usually associated with the attribution of a descriptive role to those very sentences. Statements like the following ones are impressive for their crude anti-realistic explicitness: "Really 'The proposition is either true or false' only means that it must be possible to decide for or against it. But this does not say what the ground for such a decision is like" (OC § 200); "What does this agreement [between an hypothesis and the world] consist in, if not in the fact that what is evidence in these language games speaks for our proposition?" (OC § 203); "If the true is what is grounded, then the ground is not true, not yet false" (OC § 205) (note: "the true is what is grounded", a clear-cut anti-realistic slogan). To put it in a nutshell, the truth predicate does not correspond to a guiding principle of assertoric discourse different from warranted assertibility (this is the main conclusion—needless to say, an highly debatable one—drawn by Wittgenstein from the Disquotational Schema). Furthermore, assertibility, truth-aptness, belonging to the doxastic and epistemic sphere, being an object of discovery and not of invention, are features of the contents of declarative sentences which are always bound together, in opposition to the characteristics of the contents of sentences which express norms or rules of grammar.

In discussing the purported correspondence to reality of the sentences of pure mathematics, a discussion prompted by Hardy's article *Mathematical Proof*, a sharp distinction is traced between the sense that the expression "correspondence to reality" has in the case of empirical sentences and the sense it can be given in the case of those of pure mathematics: in the latter case, the correspondence with reality is at most a matter of the relation that a rule entertains with the social practice of its use[3]. If a radically anti-realist interpretation of the grammar of a certain class of sentences is marked by the rejection of the idea that they have an assertoric content, that they "convey thoughts", then Wittgenstein's conception of pure mathematics is radically anti-realist (and the continuity with the *Tractatus* 6.21 thesis that "A proposition of mathematics does not express a thought" is really striking).

What results from my reading is that two distinct kinds of anti-realism are to be found in Wittgenstein's philosophy: a mild anti-realist stance based

3. See Hardy 1929 and LFM, Lectures 25 and 26.

on the identification of truth with warranted assertibility, whose field of application is constituted by empirical sentences, the sentences belonging to the doxastic and epistemic sphere; and a radical anti-realism, understood as the denial of the possession of an assertoric content, as regards the sentences of the critical classes, and of mathematics in particular. However, we cannot rest content with this conclusion: the question inevitably arises of how radical anti-realism can coexist with the vital role mathematical proof has in Wittgenstein's conception, and the answer to it is anything but easy to find. As a matter of fact, a large portion of Wittgenstein's writings on mathematics, especially those of the intermediate phase, is devoted to a thorough discussion of the theme of the relation between truth and provability. As we know, radical anti-realism would forbid speaking of the truth of mathematical sentences, once they have been deprived of any assertoric content. Nonetheless, in the texts of the years 1929–1933, Wittgenstein often points to provability as the only legitimate *Ersatz* of truth in the case of mathematical sentences. To be accurate, things are even more intricate because Wittgenstein, in those writings, strives to make room for a notion of a mathematical proposition which is analogous to that of an empirical proposition, where the analogy is founded on verificationist grounds, that is, on the existence of a general method of decision (an algorithm) for the sentences belonging to a whole system (the system of arithmetical identities, for instance). Beyond those narrow confines, that is, in the case of the so-called isolated propositions, Wittgenstein's prevailing conception is that it is only through the construction of a proof that a mathematical sentence gets a definite sense. It seems to me that it is at this juncture that the answer to the question we are dwelling on can be found. As a consequence of the view of proof as a means not to establish the truth of a mathematical sentence but to determine its sense, truth-aptness is ruled out and, at the same time, as in many other variants of anti-realism, provability replaces truth as the key-notion. In my opinion, the peculiarity of Wittgenstein's philosophy of mathematics lies precisely in the way in which a central role is preserved for proofs within a conception that denies truth-aptness to sentences of pure mathematics.

Let us see in outline how the relation between a theorem and its proof is to be conceived from Wittgenstein's standpoint. Even though his approach is strongly conditioned by his adhesion to the linguistic turn, I will be speaking freely of concepts and conceptual connections where Wittgenstein would have spoken of meanings and rules of grammar. First of all, the distinction between propositions belonging to a system and isolated propositions, which is so important in the writings of the intermediate phase, leads us back to the distinction between verification and proof proper which is pivotal in the great anti-logicist tradition in philosophy of mathematics, whose eminent ancestors were Descartes and Kant and whose more recent supporters are, despite notable differences between them, Poincaré, Brouwer, Polya and Lakatos. The process of verification of a proposition belonging to a system is nothing but the application of an algorithm and presupposes a static conceptual background, whereas the proof of an isolated proposition, the proof proper, brings about a conceptual dynamic, a development of concepts through the creation of new inferential connections. This is the sense in which Wittgenstein's philosophy of mathematics recalls many typical themes dealt with by heuristics, which are deployed against the "harmful invasion of mathematics by logic" (RFM V § 24) and, in my opinion, can be read as a sort of semantic inferentialism.

I want to illustrate my point by considering a celebrated case, that is, Andrew Wiles' proof of Fermat's Last Theorem (the statement: if n is any natural number greater than 2, the equation $a^n + b^n = c^n$ has no solutions in integers, all different from 0). As is well-known, a decisive step towards the proof was Gerhard Frey's conjecture (the so-called epsilon conjecture), subsequently proved by Ken Ribet, that, if a solution of the equation $a^n + b^n = c^n$ existed (where a, b, c are integers all different from 0 and n is any natural number greater than 2), then an elliptic non-modular curve would exist. It follows that, if every elliptic curve is modular, then Fermat's Last Theorem is true: but the statement that every elliptic curve is modular is the Taniyama-Shimura Conjecture, which sets up a new connection between two apparently very distant worlds, the world of elliptic equations and that of modular forms. The Taniyama-Shimura Conjecture was brilliantly proved by Andrew Wiles in 1995, and from it together with the Frey-Ribet Theorem as the major premise, Fermat's Last Theorem immediately follows by Modus Ponens. Now, let us ask ourselves: how can this story be told in Wittgensteinian terms? I would say this way: the meanings of the non-logical terms occurring in the sentence expressing Fermat's Last Theorem, together with its logical form, provide the sentence with a minimal sense only (which makes it a mere stimulus for mathematical research); this sense is quite different from the one the sentence acquires when it is immersed in a definitely more complex network of deductive connections, such as the one established by the Frey-Ribet Theorem, linking it to the theory of elliptic equations and modular forms, with its highly sophisticated concepts and techniques of proof. From a semantic point of view, Wittgenstein's could be aptly described as a sort of inferentialistic and constructivistic approach: the sense of a mathematical sentence changes whenever its deductive connections with other sentences, which are created through the construction of communally accepted proofs, modify its position in the overall network of mathematical definitions, conjectures and theorems.

At this point, a new question begs to be answered: if, for Wittgenstein, speaking of conceptual connections is simply a way of speaking of grammar rules, should the objectivity of the theorems of pure mathematics be ultimately rooted in the agreement in the reactions of all those who are faced with the inferential steps in their proofs? (in the sense that they all agree, or otherwise, in taking them as correct – an agreement in acting, rather than in believing). What I have said so far are just prolegomena to any enquiry that aims at giving a reasonable answer to this difficult question and at supplying a faithful reconstruction of the picture of mathematics Wittgenstein offers as an alternative to the traditional philosophical views and to their everlasting conflicts.

References

Blackburn, Simon 1990: "Wittgenstein's Irrealism". In: Rudolf Haller & Johannes Brandl (eds.), *Wittgenstein: Towards a Re-Evaluation*. Wien: Hölder-Pichler-Tempsky, 13–26. Reprinted in Blackburn 2010, 210–219.

— 1998: "Wittgenstein, Wright, Rorty and Minimalism". *Mind* 107, 157-181.

— 2010: *Practical Tortoise Raising and Other Philosophical Essays*. Oxford: Oxford University Press.

Conant, James 1997: "On Wittgenstein's Philosophy of Mathematics". *Proceedings of the Aristotelian Society* 97, 195–222.

Hardy, Godfrey Harold 1929: "Mathematical Proof". *Mind* XXXVIII, 149, 1–25.

Lovibond, Sabina 1983: *Realism and Imagination in Ethics.* Oxford: Blackwell.

Monk, Ray 2007: "Bourgeois, Bolshevist or Anarchist? The Reception of Wittgenstein's Philosophy of Mathematics". In: Guy Kahane, Edward Kanterian & Oskari Kuusela (eds.), *Wittgenstein and His Interpreters*. Oxford: Blackwell Publishers, 269–294.

Putnam, Hilary 1996: "On Wittgenstein's Philosophy of Mathematics". *Proceedings of the Aristotelian Society* 70, 243–264.

Putnam, Hilary 2001: "Was Wittgenstein Really an Anti-Realist about Mathematics?". In: Timothy McCarthy & Sean C. Stidd (eds.), *Wittgenstein in America*. Oxford: Oxford University Press, 140–194.

Putnam, Hilary 2008: "Wittgenstein and Realism". *International Journal of Philosophical Studies* 16 (1), 3–16.

Wright, Crispin 1992: *Truth and Objectivity*. Cambridge, MA: Harvard University Press.

Wright, Crispin 2003: *Saving the Differences. Essays on Themes from Truth and Objectivity*. Cambridge, MA: Harvard University Press.

Grazer Philosophische Studien
89 (2014), 23–38.

MATHEMATICAL PROPOSITIONS AS RULES OF GRAMMAR

Severin SCHROEDER
University of Reading

Summary
There is a tension between Wittgenstein's claim that mathematical propositions are rules and his emphasis on the practical usefulness of mathematics. As Wittgenstein himself asks: 'How can the mere transformation of an expression be of practical consequence?' (RFM 357) I argue that Wittgenstein's answer is that mathematical propositions are not merely stipulative definitional rules: they forge connections between independently comprehensible concepts. A proposition expressing an empirical correlation may be fixed as a mathematical rule and thus be made independent of experience. I then consider and respond to further objections to Wittgenstein's position.

1. Are mathematical propositions descriptions of timeless abstract entities (Platonism) or are they generalisations of empirical observations (Mill; Formalism)? Neither, according to Wittgenstein. Platonism and empiricism share the assumption that mathematical propositions are *descriptions* of something, which is exactly what Wittgenstein rejects. They are not descriptions, but norms of representation: rules of grammar. That explains their peculiar dignity: their certainty and necessity. The mathematical reliability and inexorability is ultimately our own reliability and inexorability in insisting on those rules and not allowing any exceptions to them.

2. There is a fairly uncontroversial sense in which some mathematical propositions can be called 'rules'.

First, there are definitions. 1 + 1 = 2, 2 + 1 = 3, 3 + 1 = 4, etc. serve to define both the series of natural numbers and the operation of addition. These are axioms, or basic rules of the arithmetical calculus.

Secondly, there are non-basic, but simple equations, such as in the times tables, which we memorize at an early age and apply when doing longer calculations.[1] Thus, '$3 \times 9 = 27$' is applied as a rule when working out the long multiplication: 399×39 (cf. PLP 53).

Thirdly, at a slightly more advanced level there are algebraic formulae that are both proven true and, afterwards, memorised or consulted for repeated application. E.g., the cosine rule or the quadratic formula.

However, such cases are *not* what Wittgenstein has in mind when he calls mathematical propositions rules. 'If one says the mathematical proposition is a rule,' he writes, 'then of course not a rule in mathematics' (MS 127, 236).[2] Rather, on his view, they are rules of grammar, and, what is more, not the grammar of mathematics, but the grammar of non-mathematical language. The idea goes back all the way to the *Tractatus*, where mathematics is characterised by the following three propositions:

> Mathematical propositions are equations. (TLP 6.2)
>
> If two expressions are combined by the sign of equality, that means that they can be substituted for one another. (TLP 6.23)
>
> … in real life … we make use of mathematical propositions only in inferences from propositions that do not belong to mathematics to others that likewise do not belong to mathematics. (TLP 6.211)

Thus, the equation '$2 + 3 = 5$' is a grammatical rule for the use of number words in a natural language, licensing, for instance, the inference from 'I have two coins in my right pocket and three coins in my left pocket' to 'I have five coins in my pockets'.

Wittgenstein lays particular stress on the dependence of mathematics on its having applications *outside* mathematics. That is what turns a mere calculus, a game of manipulating signs according to certain rules, into mathematics.

> it is mathematics, I should think, when it is used for the transition from one proposition to another. (BT 533)
>
> It must be essential to mathematics that it can be applied. (BT 566)
>
> I want to say: it is essential to mathematics that its signs are also employed in *mufti*.
>
> It is the use outside mathematics, that makes the sign-game into mathematics. (RFM 257; cf. 295)

1. In LWL Wittgenstein calls them 'definitions'.
2. *Wenn man sagt, der mathematische Satz ist eine Regel, so natürlich nicht eine Regel in der Mathematik*. [MS 127, 236; post 4.3.44]

> mathematical propositions containing a certain symbol are rules for the use of that symbol, and … these symbols can then be used in non-mathematical statements. (LFM 33; cf. 256)

3. Equations, according to Wittgenstein, are rules for connecting concepts, thus forging a new enriched concept (RFM 412, 432). The grammatical rule this provides is that where one side of the equation applies, the other one must apply too: If, for instance, something is 2 + 3, it must also be 5.

In this way, mathematical propositions are said to fix sense, not to establish any substantive truth:

> For the mathematical proposition is to show us what it makes SENSE to say. (RFM 164b)

> If you know a mathematical proposition, that's not to say you yet know *anything*.
>
> I.e., the mathematical proposition is only supposed to supply a framework for a description. (RFM 356)

This is provocative, and Wittgenstein himself at times felt provoked by it. He acknowledged that there appears to be a tension between this claim — what might be called a non-cognitivist account of mathematics — and mathematics' well-known prognostic potential and practical usefulness. Equations are to be mere transformations of expressions, but 'How can the mere transformation of an expression be of practical consequence?' (RFM 357)

> I can use the proposition '12 inches = 1 foot' to make a prediction; namely that twelve inch-long pieces of wood laid end to end will turn out to be of the same length as one piece measured in a different way. Thus the point of that rule is, e.g., that it can be used to make certain predictions. Does it lose the character of a *rule* on that account? (RFM 356; cf. 381a)[3]

Very often I can calculate what will happen: mathematics teaches me an observable result. An area of 7 by 5 foot is to be covered with tiles each one foot square. How many tiles are required? An elementary calculation tells me that I shall need 35 tiles. So it appears that maths can serve to discover an empirical truth, and not just to fix sense.

---

3. Cf. MS 163, 62r: *'Die Mathematik eine Grammatik?? Aber sie hilft uns doch Vorhersagen machen!' – Sie* hilft *uns*.

Moreover, if an equation were just a rule to determine what makes sense and what doesn't, the application of a miscalculation should result in nonsense. For example:

(T) In order to cover an area of 7 by 5 foot I need 37 tiles of one foot square.

should make no sense; but it seems much more natural to say that it's just false. After all, I can wonder if it might not be true (RFM 62: I-67); I can try and convince myself empirically that I don't need 37 tiles. And what's more, it is not inconceivable that (T) *might* be true. We could perhaps imagine that somehow when we put down 35 tiles there still remain two empty squares; and, strangely, when we count the laid out tiles we always get 37 (cf. RFM 91b).

4. In the early 1930s, Wittgenstein (as reported by Waismann) denied that a mathematical proposition allows us to predict empirical observations:

> It may seem as if the equation 5 + 7 = 12 entitles us to make statements about the future, namely to predict what number of shillings I shall find if I count the ones I have in each pocket [5 and 7]. But this is not the case. Such a statement about the future is justified by a physical hypothesis which stands outside the calculus. If a shilling suddenly disappeared, or if a new one suddenly came into existence while we were counting, we should not say that experience had disproved the equation 5 + 7 = 12; similarly, we should not say that experience had confirmed the equation. (PLP 51f.)

On this view, mathematics only connects the concepts of *5 and 7* and of *12*, forming and insisting on a joint concept, according to which whatever falls under the one falls under the other. Using the joint concept to move from one description to the other does not involve any physical hypothesis. If you have 5 and 7 shillings in your pockets, then *ipso facto* you have 12 shillings in your pockets. But things are different if we talk about predicting what *will* be found in your pockets the next moment. The mathematical substitution rule '5 + 7 = 12' alone cannot guarantee that having counted 5 and 7 shillings *now*, I shall count 12 shillings a moment later. That prediction also involves the physical hypothesis, or assumption, that coins have a certain durability: that they don't suddenly disappear, coalesce or multiply.

One might reply, however, that although the durability of coins must of course be presupposed, the fact remains that the prediction in question is arrived at and justified by the equation. When asked: ‘What makes you think that you’ll find 12 shillings in your pockets’, the answer is not: ‘What I know about the physical nature of coins’, but rather: ‘I counted 5 in one pocket and 7 in the other, and 5 + 7 = 12’.

5. The question remains: how can Wittgenstein’s view of mathematical propositions as rules of grammar be reconciled with their prognostic usefulness?

There is one striking remark in Wittgenstein’s later writings which almost sounds as if he retracted his position in light of mathematics’ role in predictions. He writes:

> It is clear that mathematics as a technique for transforming signs for the purpose of prediction has nothing to do with grammar. (RFM 234c)

For the moment, however, I shall put this puzzling remark to one side, and return to it only at the end of my paper.

6. In his later writings Wittgenstein does not say categorically that mathematical propositions are rules of grammar, but expresses himself more cautiously:

> I have no right to want you to say that mathematical propositions are rules of grammar. I only have the right to say to you, “Investigate whether mathematical propositions are not rules of expression, paradigms — propositions dependent on experience but made independent of it …” (LFM 55)

> There is no doubt at all that *in certain language-games* mathematical propositions play the part of rules of description, as opposed to descriptive propositions.
>
> But that is not to say that this contrast does not shade off in all directions. (RFM 363)

> What I am saying comes to this, that mathematics is *normative*. (RFM 425)

> Mathematical propositions are essentially akin to rules … (RPP I § 266)

These passages suggest that although Wittgenstein insisted on the essentially normative, and hence rule-like character of mathematical propositions, he was also alive to differences between mathematical propositions and ordinary grammatical propositions, such as 'A bachelor is an unmarried man'. One crucial difference is mentioned in the passage from the 1939 lectures: unlike ordinary grammatical propositions, mathematical propositions are in a peculiar way 'dependent on experience' — before they're made independent of it. In a sense they can be said to originate from and reflect experience.

7. Our arithmetic is based on experiences with counting objects, experiments such as this one:

> Put two apples on a bare table, see that no one comes near them and nothing shakes the table; now put another two apples on the table; now count the apples that are there. You have made an experiment; the result of the counting is probably 4. (RFM 51)

Because this kind of experiment leads reliably to the same result, we introduce a corresponding rule, which turns, so to speak, a highly probable outcome into a conceptual necessity.

> I believe that it will probably always be so (perhaps experience has taught me this), and that is why I am willing to accept the rule: I will say that a group is of the form *A* [5 strokes] if and only if it can be split up into two groups like *B* [2 strokes] and *C* [3 strokes]. (RFM 62)

Thus, after first finding the concepts of *2 and 3* and of *5 empirically* correlated, we come to introduce '2 + 3 = 5' as a *mathematical* proposition, that is: a norm of representation. If now the original experiment leads to a different result, we shan't accept it: We shall insist that we must have made a mistake or that something strange must have happened to account for this deviation from our norm.

This was an extremely simple example. In other cases, when first we approach a question of adding larger numbers of objects together, let alone multiplying them, the outcome may not be immediately obvious. In some cases the outcome may even be surprising, contrary to what we expected (RFM 63b).

8. Not only are elementary mathematical propositions based, genetically, upon corresponding empirical propositions, or experiences, they also

require that our experiences continue to be, by and large, in agreement with our calculations. Although no individual experience can disprove an arithmetical equation, used as a norm of representation, a regular discrepancy between rule and experience would undermine the rule's usefulness and eventually make us abandon or change it.

> This is how our children learn sums; … one makes them put down three beans and then another three beans and then count what is there. If the result at one time were 5, at another 7 (say because, *as we should now say*, one sometimes got added, and one sometimes vanished of itself), then the first thing we said would be that beans were no good for teaching sums. But if the same thing happened with sticks, fingers, lines and most other things, that would be the end of all sums.
>
> "But shouldn't we then still have 2 + 2 = 4?"—This sentence would have become unusable. (RFM 51f.)

This is a central aspect of Wittgenstein's account of mathematics that is well worth emphasising. An equation, such as '2 + 2 = 4', is not an empirical generalisation, and hence no contrary experience can disprove it. On the other hand, it is not entirely independent of experience either. It is essentially a norm for describing countable things, like beans and sticks, and hence dependent on its suitability for the purpose (RFM 357c).

9. In his 1939 lectures Wittgenstein points out emphatically and repeatedly that corresponding to an arithmetical equation there is an empirical statement expressed in similar or even the same words that it is important to distinguish from the mathematical proposition (LFM 111). One might be inclined here to think of a pair like:

(M) 5 + 7 = 12
(MA) 5 apples and 7 apples are 12 apples.

But as so often in his philosophy, Wittgenstein tells us not to look only at forms of words, but at the use made of them (cf. LC 2). The occurrence of an empirical term like 'apple' is no reliable indication that we are considering an empirical statement. As he remarks elsewhere, 'mathematical propositions might quite well be expressed in terms of people, houses, or what not' (LFM 116; cf. 113). A norm of representation can be taught by giving a possible instantiation. A term such as 'apple' may function somewhat like a variable, indicating that an arithmetic equation is essentially applicable to things, and not, as Platonists have it, a self-sufficient statement about abstract objects.

So, although apparently about apples, (MA) can well be used as a mathematical proposition: as an expression of a norm of representation. On the other hand, the naked equation (M) could, according to Wittgenstein, be taken as an empirical generalisation. So the distinction between the mathematical and the non-mathematical use of number sentences need not coincide with that between the two kinds of formulations, but can cut right across it. Whatever formulation you choose it can be understood either way:

> The point is that the proposition "25 × 25 = 625" may be true in two senses. If I calculate a weight with it, I can use it in two different ways.
>
> First, when used as a prediction of what something will weigh — in this case it may be true or false, and is an experiential proposition. I will call it wrong if the object in question is not found to weigh 625 grams when put in the balance.
>
> In another sense, the proposition is correct if calculation shows this — if it can be proved — if multiplication of 25 by 25 gives 625 according to certain rules.
>
> It may be correct in one way and incorrect in the other, and vice versa.
>
> It is of course in the second way that we ordinarily use the statement that 25 × 25 = 625. We make its correctness or incorrectness independent of experience. In one sense it is independent of experience, in one sense not. (LFM 41; cf. 292)

The point to note is that, on Wittgenstein's account, for all mathematical propositions there are analogous empirical statements to the effect that things will, as a matter of fact, turn out in accordance with the mathematical norm (LFM 111). For (M) '5 + 7 = 12', for instance, there is the prediction that if you count 5 apples and 7 apples then counting the total will indeed yield 12. That prediction may occasionally, rarely, be false.

10. If we now compare elementary mathematical propositions with ordinary grammatical propositions—such as:

(B) A bachelor is an unmarried man.

—it should become clear that they are significantly different. (B) is constitutive of the meaning of its subject term: it explains what the word 'bachelor' means. 'Bachelor' and 'unmarried man' are just two labels for the same concept. Hence, if you understand the expressions, you cannot ever know that one of them applies without knowing that the other one applies as well. By contrast, (as famously pointed out by Kant)[4] *7 + 5* and *12* are different concepts: they have different criteria of application (counting to 7 and counting to 5 versus counting to 12) (cf. RFM 357). Hence it is *possible* to count on a given occasion 7 and 5 objects, but only 11 altogether (or, to use Wittgenstein's example, 25 × 25, but not 625) (RFM 358e). In this case we have, initially, two distinct concepts, independently comprehensible—'Only through our arithmetic do they *become one*' (RFM 358; cf. 359a). Note the emphasis on 'become': If mathematical propositions are grammatical propositions they are essentially *additional* ones: *further* rules for terms that are already understandable without them. Mathematical propositions are *enriching existing meanings*. The norm expressed by a grammatical proposition like (B), by contrast, does not *change* or *enrich* the meaning of the word 'bachelor', it gives it its meaning in the first place.

11. Mathematical propositions are more like another type of grammatical proposition, fairly common in scientific discourse. Sometimes what used to be an empirical discovery is later made part of a definition, for example, the velocity of light or the key properties of an acid. Thus, Wittgenstein writes with reference to mathematical propositions:

> Every empirical proposition may serve as a rule if it is fixed, like a machine part, made immovable, so that now the whole representation turns around it and it becomes part of the coordinate system, independent of facts. (RFM 437)

A mathematical proposition has been grafted onto a corresponding empirical observation. By contrast, it could never have been empirically discovered that a bachelor is an unmarried man.

4. *Critique of Pure Reason* (tr. N. Kemp Smith), B 15: But if we look more closely we find that the concept of the sum of 7 and 5 contains nothing save the union of the two numbers into one, and in this no thought is being taken as to what that single number may be which combines both. The concept of 12 is by no means already thought in merely thinking this union of 7 and 5 …

12. If elementary mathematical propositions are essentially additional rules for combining existing concepts, the question is whether these rules become fully integrated in our language, as Wittgenstein seems to suggest when he calls them 'grammatical' or 'instruments of language' (RFM 99, 162, 164-6, 358d, 359a). There are, I believe, reasons to return a negative answer: reasons not to regard mathematics—except perhaps for its very rudiments—as part of the grammar of our language.

13. What characterizes a grammatical proposition is that, as it determines what makes sense, its negation, or a sentence that violates the norm it expresses, is nonsense. Is that also true of mathematical propositions?

At the most elementary level this may be so. The sentence 'I had two coffees in the morning and two in the afternoon, so I had only three overall today.' is patently inconsistent. It might well be dismissed as not only false, but nonsensical.

But suppose someone said:

(R) The pitch of the roof of my lean-to garage is 15° to the horizontal and the roof extends 5.36 meter horizontally from the wall, and one side of the roof is 1.32 meter higher than the other.

Would we dismiss *that* as nonsense? Certainly not straightaway, for as far as we know it might even be true. Only a trigonometric calculation shows that:

(A) If a right-angled triangle has an angle of 15° and the adjacent side is 5.36 then the opposite must be 1.44.

So (R) cannot be correct after all. And yet one can *believe* it to be correct —which speaks against regarding it as nonsense. For where there is no sense, there is nothing to believe.

14. Wittgenstein discusses the issue of believing false equations (RFM 76–9), and seems to suggest that when I mistakenly believe that 16 × 16 = 169, I do not believe a mathematical proposition; rather, I mistakenly believe that '16 × 16 = 169' *is* a mathematical proposition, a rule of our mathematical grammar—which it isn't. It has no meaning *in arithme-*

tic, just as moving a pawn backwards is not a move in chess, not even a bad one.

However, our question was not how one could believe a false *mathematical* proposition, but how one could believe a false *non-mathematical* proposition such as (R) that is in conflict with a mathematical norm of representation (A) and should be ruled out by it. Of course, as remarked earlier, it is not impossible for a trigonometric norm of representation to be put in terms of the shape of the roof of a lean-to garage, but that is hardly a natural understanding of (R). The corresponding trigonometric norm would more naturally be put in the form of a conditional, like (A). It is much more likely that we take (R) as an empirical statement: as the speaker's report of his measurements. I can certainly believe that those are the correct measurements *before* I've done the maths. Afterwards I shall think that the speaker must have made a mistake (or that the roof isn't straight, so that one cannot really speak of a pitch of 15°). Yet the fact remains that one can understand (R) as an empirical statement and believe it to be true.

15. That impression is reinforced by Wittgenstein's considerations about the relation between a mathematical proposition and corresponding empirical observations, expounded above.

> The rule doesn't express an empirical connection but we make it because there is an empirical connection. (LFM 292)

The rule's usefulness depends on its continued empirical appropriateness. It is not only that we reject (R) in the light of (A): that we insist that some of the measurements of the roof must have been inaccurate. It is also that when, in such a case, we measure or count again with greater care we shall almost certainly find our empirical observations in agreement with the rule. In this case: if the other measurements prove reasonably accurate, we shall find that

(R1) one side of the roof is indeed about 1.44 meter higher than the other side.

That is to say, we take an empirical proposition, such as this, (R1), as confirmation of a mathematical proposition, such as (A); confirmation not of the *truth* of (A)—for (A) is a rule, not a generalisation—, but confirmation of its suitability and usefulness in the light of experience. But to take (R1) as an *empirical* proposition means to envisage the possibility of its being true *or false*; which means to envisage that something like (R) might have been true. So (R), although ruled out by a mathematical proposition (A), cannot be nonsense (as it would have to be if a mathematical proposition were not only a rule, but a *grammatical* rule: a norm for what makes sense in our language).

16. If, then, mathematical propositions cannot be regarded as grammatical rules, how else are we to characterise them? We can still agree with Wittgenstein that mathematical propositions are norms of representation. For such norms to be in force and to be rigorously insisted on they need not be linguistic norms. They may even be called 'grammatical' if that is taken to refer not to the grammar of our language, but only to a specific form of discourse, or, taking the word 'grammar' in a figurative sense, we may speak of the 'grammar', i.e. the system of rules, of a certain set of activities or of some institutionalised form of life. In a laconic remark in *Philosophical Investigations* Wittgenstein suggests that theology can be regarded as grammar (PI § 373), providing rules for what can be said meaningfully about God. But these rules are binding only within a certain religious community. Thus, for a believer God is by definition omnipotent and benevolent.[5] To question these attributes doesn't make any sense within religious discourse: it would be 'ridiculous or blasphemous' (AWL 32) . And yet an agnostic or atheist may well do so. You can step outside religious language, flouting its grammatical norms, while remaining within language.

17. Consider the following augury language-game: People have a sophisticated calculus for determining propitious days for travelling. The parameters are the number of people travelling together, their average age, and the distance to be travelled. A certain algorithm results in three numbers between 1 and 31, which specify propitious days of the month for setting off.

N * A * D = (x, y, z)

5. Z § 717: '"You can't hear God speak to someone else, you can hear him only if you are being addressed".—That is a grammatical remark.

Thus, if 3 people aged 25, 30 and 35 years respectively want to make a journey of 217 miles, they might carry out calculations resulting in the following formula:

(AC) $3 * 30 * 217 = (17, 18, 23)$.

This means that the 17th, the 18th and the 23rd would be propitious days for undertaking their journey.

Such an augury calculus would, presumably, be a kind of mathematics.[6] For its practitioners, (AC) serves as a 'grammatical' rule, endorsing certain statements and ruling out others. The following, for example, would be ruled out by the formula as 'ungrammatical':

(J) The 5th May would be a propitious day for 3 people aged 25, 30, and 35 years respectively to make a journey of 217 miles.

And yet, clearly, (J) is not linguistically flawed. It might even be true. (Note that for the augury calculus to have a real application, 'propitious day' must not be defined by the calculus.)

18. These are examples of rules of 'grammar' in a wider sense of the word: not the grammar of a language like English or German, but the 'grammar' (so to speak) of a certain kind of discourse or a certain kind of prognostic activity. My suggestion is that if we follow Wittgenstein in regarding mathematical propositions as grammatical norms, we need to understand the word 'grammatical' in a similar way: not as determining what makes sense in a natural language, but rather fixing sense and non-sense in a specific kind of discourse or activity. That is, roughly speaking, an activity and discourse in which we try to develop and apply a system of *calculating* quantities, rather than simply counting or measuring them.[7]

At this point it may also be apposite to mention that Wittgenstein repeatedly suggests that calculations need not be laid down in sentences:

---

6. But cf. RFM 399c.

7. Note that my comparison between mathematics and theology and augury is concerned only with the way the normativity of each of these activities is restricted and does not coincide with linguistic normativity. I am not suggesting that mathematics is only a matter of faith or superstition.

that mathematics is primarily an activity, and not necessarily an entirely linguistic activity (RFM 93d). That, too, would suggest that the norms set up by mathematics are not so much linguistic norms, as methodological rules, roughly speaking, for dealing with quantities.

> A proof is an instrument—but why do I say "an instrument of language"? Is a calculation necessarily an instrument of language, then? (RFM 168)

> Mathematics consists of calculations, not of propositions. (MS 121, 71v)[8]

> The grammatical rules are comparable to rules about the procedures in measuring periods of time, distances, temperatures, forces, etc. etc.. Or: these methodological rules are themselves examples of grammatical rules. (MS 117, 138f.)[9]

19. Finally, I'd like briefly to return to the remark quoted in §5 above, which seems to be in stark conflict with Wittgenstein's key idea of mathematics as grammar:

> It is clear that mathematics as a technique for transforming signs for the purpose of prediction has nothing to do with grammar. (RFM 234c)

[i] Here is one possible interpretation: Perhaps what Wittgenstein means here by 'a technique for transforming signs' are operations that—unlike addition, subtraction, multiplication or division—have no immediate application to non-mathematical sentences. For remember, what Wittgenstein has in mind when he calls mathematical propositions grammatical rules are rules for non-mathematical language; especially rules for 'inferences from propositions that do not belong to mathematics to others that likewise do not belong to mathematics' (TLP 6.211). Yet large parts of the techniques of, say, trigonometry or differentiation are hardly applicable to non-mathematical sentences. The beginning and end result of such transformations may be expressible in English, and their connexion may serve as a norm for connecting sentences in English, but the intermediary steps are moves in a calculus that have no immediate bearing on what English propositions are acceptable.

---

8. *Die Mathematik besteht aus Rechnungen, nicht aus Sätzen.*

9. *Die grammatischen Regeln sind zu vergleichen Regeln über das Vorgehn beim Messen von Zeiträumen, von Entfernungen, Temperaturen, Kräften, etc. etc. Oder auch: diese methodologischen Regeln sind selbst Beispiele grammatischer Regeln.*

The cosine rule ($a^2 = b^2 + c^2 - 2bc \cos A$), for example, is an equation that does not fit the *Tractatus* model for the use of mathematical equations: as a rule licensing the transition from one non-mathematical proposition to another. For *cosine* is an exclusively mathematical concept that does not feature in any non-mathematical proposition (about the size and shape of a pitched roof, for instance). Hence such an equation cannot serve as a grammatical proposition in the way an arithmetical sum does. Rather, it plays a role only inside mathematics, transforming one mathematical proposition into another one.

However, the ultimate aim of such inner-mathematical transformations can still be a correct empirical statement ('for the purpose of prediction'), so the wording 'has nothing to do with grammar' would be an exaggeration: Indirectly such inner-mathematical transformations could still be said to have something to do with grammar. Therefore this interpretation is not entirely satisfactory.

[ii] Perhaps a more plausible explanation of what Wittgenstein means here is this: If on the basis of a mathematical calculation I predict that my journey from Charlbury to Heathrow will take 2 hours 50 minutes, but then the train is late and in fact it takes me 3 and a half hours to get to Heathrow, I shall not insist on the result of my calculation. In this case, I shall *not* use the calculation as a grammatical norm, checking and correction empirical observations.

Again, I calculated that in order to cover an area of 7 by 5 foot, I should need 35 square foot tiles. If counting the tiles I put down yielded 36, I should say that I must have miscounted. I should count again until I got a result that agreed with my calculation. In this case the multiplication serves as a grammatical norm to control and correct my counting. If, on the other hand, I broke 2 tiles when trying to fit them in,—my observation that in the end I needed 37 tiles for the job, would not be revised and brought in agreement with my calculated prediction that I'd need only 35 tiles.

That is to say, there are two kinds of cases in which a prediction based on a calculation can be found to disagree with the observed outcome. It may be that there are additional factors not taken into account in the calculation. Or it may be that the observation or measurement was inaccurate. It is only in what we take to be the latter case that calculations serve as norms of representation, as checks on our observations. Perhaps the remark under discussion (RFM 234c) is supposed to be a reminder of the former case in which calculated predictions play no normative role, but are regarded as mistaken.[10]

10. I am grateful to David Dolby for his comments on an earlier version of this paper.

Grazer Philosophische Studien
89 (2014), 39–59.

HOW ARITHMETIC IS ABOUT NUMBERS. A WITTGENSTEINIAN PERSPECTIVE

Felix MÜHLHÖLZER
Georg-August-University, Göttingen

Summary
When the aboutness of first-order arithmetic is given a precise formulation by using the model-theoretic notion of *interpretation*, it cannot be caught uniquely: there are always non-standard models. It is possible to single out the intended standard model in the meta-language, but then we *refer* to this model. From a Wittgensteinian perspective, reference is categorially different from interpretation because it essentially involves the use of our signs; interpretation, on the other hand, is of a purely mathematical nature, totally abstracted from use. In this way the seeming tension between these two ways of seeing the aboutness of arithmetic dissolves.

„Verstehen" und „meinen" sind Worte wie alle anderen.
(MS 116, 3)

1. *Philosophical Investigations,* §10

The very first section of the *Philosophical Investigations* may give the impression that Wittgenstein deprives numerals of reference. After having presented what is often called the "Augustinian picture" of language, with its fundamental idea that every word has a meaning and that this meaning is the object for which the word stands, Wittgenstein, with his example of the greengrocer who sells apples, tells an alternative story in which the words of our language appear as something we *operate* with and not as something *standing for* certain objects. At the end of §1 of PI he considers the number-word "five" and writes: "But what is the meaning of the word »five«?—No such thing was in question here, only how the word »five« is used."

In his middle period we in fact find passages where Wittgenstein seems to say with respect to the mathematical vocabulary in general that it does not refer to anything. So in the *Big Typescript*:

> Mathematics consists entirely of calculations.
>
> In mathematics *everything* is algorithm, *nothing* meaning [*nichts* Bedeutung]; even when it seems there's meaning, because we appear to be speaking *about* mathematical things in *words*. What we're really doing in that case is simply constructing an algorithm with those words. (BT[1], 748f.)

Even more explicit is the following statement: "[M]athematics is a calculus and therefore is really about nothing" (BT, 532). General claims of this sort are remnants of the dogmatic thinking of the *Tractatus*. Wittgenstein has the idea that the essence of mathematics is calculation, and he draws conclusions from this claim; for example, that there cannot be a metamathematics in Hilbert's sense, that is, a mathematics that is *about* formalized theories.[2] Mathematics simply isn't 'about' anything, according to the dogmatic point of view which is still present in the *Big Typescript*.

Admittedly, algorithmic procedures are something very specific and characteristic, and it is a tempting idea to see in them the essence of mathematics.[3] But our real mathematical practice, including our way of speaking within this practice, pays no attention to such philosophical prejudices, and the post-*Tractatus* Wittgenstein tries to be sensitive to our practice as it is. Thus, in the *Big Typescript*, contrary to its dogmatic assertions, when having raised the question: "What are numbers?", Wittgenstein immediately gives the answer: "The meanings of numerals [Die Bedeutungen der Zahlzeichen]" (BT, 569). This can only mean: what the numerals *refer to*. And in his *Lectures on the Foundations of Mathematics* of 1939 he remarks: "One might say that [the proposition »20 + 15 = 35«] is a statement about numbers. Is it wrong to say that? Not at all; that is what we call a statement about numbers." (LFM, 112) In these *Lectures* he then avoids the contradiction of the *Big Typescript*, the contradiction

1. Cited according to the original pagination in TS 213.

2. This statement, which *prima facie* may sound incomprehensible, is discussed and defended in Mühlhölzer 2012 by allowing mathematical terms to 'refer' in a way that is consistent with the subsequent considerations of the present paper.

3. Astonishingly enough, Bourbaki regards calculation as the essence of *algebra*; see the first sentence of the Introduction to Bourbaki 1942: "Faire de l'Algèbre, c'est essentiellement *calculer*". Without doubt, what Bourbaki means by "calculer" is very different from what Wittgenstein had in mind when using the words "algorithm" and "calculus".

between saying that mathematics isn't about anything *and* that numbers are the meanings of numerals, by drawing a distinction between two senses of the word "about", a more deflationist one, as it were, for mathematical propositions, and a more substantive one for empirical propositions. Compared with the way empirical propositions are about a certain subject matter, mathematical propositions appear very different, in such a way that one in fact might be tempted to dispute any 'aboutness' in their case. But this would contradict our normal and normally useful way of speaking, and we would do better simply to distinguish between the different ways in which these two sorts of propositions are 'about' this or that.[4] So we should simply accept that mathematics is 'about' a subject matter, and arithmetic, in particular, about numbers, and in the present paper I will mainly deal with the sense of "about" in the latter case.

In what follows, I will be oriented exclusively towards Wittgenstein's mature philosophy as presented in the *Philosophical Investigations* (and when referring to Wittgenstein, from now on I always mean the later, the mature Wittgenstein). To interpret the beginning of the *Investigations* as depriving numerals of reference would be a misunderstanding. What Wittgenstein really wants to convey is the thought that what we call "reference"—or "signification", "designation", "aboutness", and so on—is *secondary* in comparison with use. To understand the phenomenon of linguistic meaning we should be concerned with our use of words—and their so-called 'reference' has to be taken account of only afterwards. This is the point of the important §10 of PI, in which Wittgenstein stresses that what words signify should show itself "in the kind of use they have", and that our uniform talk of the form: »This word signifies *that*«, should not blind us to the deep dissimilarities in our use of different sorts of words.[5] Wittgenstein there explicitly mentions the difference between numerals on the one hand and words like »block«, »slab« and »pillar« on the other. In §10 he actually has a very restricted language game in mind, the language game described in §8, but his remark applies to all language games; and it concerns, of course, not only the expression "to signify" but also expressions

4. As Wittgenstein says, in their case "'being about' means two entirely different things" (LFM, 251). Cora Diamond and Peter Hacker emphasize the importance of this point (see Diamond 1996, 231–236; Baker/Hacker 2009, 283), but they do not explain it. In the present paper, as already in Mühlhölzer 2012, I try to clarify what can be meant by arithmetic's 'being about numbers'.

5. This section is already to be found in the so-called Urfassung of the *Philosophical Investigations*, essentially written at the end of 1936; see MS 142, 8f.; PUKG, 64.

like "to be about" and cognate ones. I read §10 as an invitation to clarify the actual role of such expressions, including their 'Witz', their 'point', as Wittgenstein tends to say. With respect to the numerals »a«, »b«, »c«, etc., which are introduced in §8, he writes in particular:

> [O]ne may say that the signs »a«, »b«, etc. signify numbers: when, for example, this removes the misunderstanding that »a«, »b«, »c« play the part actually played in the language by »block«, »slab«, »pillar«. And one may also say that »c« qualifies this number and not that one; if, for example, this serves to explain that the letters are to be used in the order a, b, c, d, etc., and not in the order a, b, d, c.

One need not stop at this point, of course, and can go on to raise the general question as to the function of our talking about signification, aboutness etc., depending on the language game one is playing. §10 of the *Investigations* suggests this question.

In the German version of this section, Wittgenstein uses the word "bezeichnen", which is a very ordinary word in German. It has been translated as "signify", which may not be as ordinary as the German word but shares with it the important property of not being charged with philosophical ballast like the word "refer", for example.[6] This is in fact very important for an adequate understanding of §10. In the case of the words "understanding" and "meaning", Wittgenstein explicitly says in MS 116, 3: "»understanding« and »meaning« are words like all others"[7], and the same he could have said about a word like "signify" ("bezeichnen"). Wittgenstein does not present a philosophical theory about 'signification'—like the many philosophical theories that exist about 'reference'—but is interested in our actual non-philosophical use of this term.

Nevertheless, in what follows I will stick to the term "refer", and to "reference", in order to connect Wittgenstein's view with the way philosophical discussions are usually conducted today. And PI §10 then embodies a striking downgrading of "reference" and its cognates, as if their main function were to remove misunderstandings, or to achieve similarly humble things. I read it as being slightly ironic and provocative, as if to say: in no case should reference be considered as primary over use; it's the other way

6. I am grateful to Severin Schroeder and Joachim Schulte for discussions about this issue. The word "signify" occurs in the original translation of PI §10 by Elizabeth Anscombe and has been adopted also by Peter Hacker and Joachim Schulte in their new translation.

7. Which, in its original German version, is the motto of the present paper; translations of manuscript passages are my own.

round. Irony, however, is certainly not enough, because questions about reference abound in philosophy, and we must keep an eye on what, from case to case, the function of talking about reference precisely *is*. In the case of sensation words, for example, Wittgenstein himself, in §244 of the *Investigations*, raises the question: "How do words *refer* to sensations? […] how is the connection between the name and the thing named set up?" I need not state Wittgenstein's way of answering this question here. Its essence is: let us look at the use of our sensation words and how this use is learnt—and more need not be brought into play. And the same must be said, now, of the words and expressions of arithmetic, like the numerals or general terms like "natural number" or "prime number". The use of the numerals mentioned in §10 is only: to remove certain misunderstandings; but, of course, other sorts of use can be easily stated. So, for example, we treat different systems of numerals—the tally notation; the positional notation, with the possibility of using different bases; or the numerals of the Romans—as systems referring to the *same* numbers. And so on. It would be pointless to try to give a survey of all possible uses of arithmetical terms. What Wittgenstein himself does is to raise certain philosophical problems, and he then considers the use of relevant terms in order to solve, or dissolve, them.

2. *The aboutness dilemma*

In this vein, I will now concentrate on one specific problem that has bothered many people, not only philosophers like Dummett and Putnam, for example, but also logicians like Skolem and Bernays, and even mathematicians like Yuri Manin.[8] It is a problem that concerns not so much specific uses of mathematical terms, but rather the question of whether terms are used at all.

The problem accrues from elementary results in model theory to the effect that many formalized theories, typically first-order theories, have nonstandard models. Perhaps the most famous case is first-order set theory which allows countable models although one can prove in it the

8. See Dummett 1963; Putnam 1980; Manin 1977, 69 (which will be discussed below). The debate between Skolem and Bernays about this problem is set out in Mühlhölzer forthcoming (relying on the description in Bellotti 2006, 182f.), where the problem is specifically expounded in the context of structuralist positions in the philosophy of mathematics, drawing on the careful discussion in Halbach/Horsten 2005.

uncountability of the set of real numbers (for example). This is the so-called "Skolem paradox" which, however, will not be my subject here. I'm interested in the emergence of nonstandard models of mathematical theories in general. In order to be sufficiently focussed, I will in this paper mainly consider first-order Peano arithmetic and its models. If not otherwise explained, the term "arithmetic" will always mean "first-order Peano arithmetic". The so-called *upward Löwenheim-Skolem theorem* alone then already shows that arithmetic has nonstandard models: there are models with arbitrary cardinality beyond the countable ones. But one can very easily show that there are also countable nonstandard models of arithmetic. One only needs to add one more constant, say "c", to Peano arithmetic and the following enumerable list of new axioms: "$0 < c$", "$1 < c$", "$2 < c$", and so on for all numerals. Then any finite set of axioms of this new theory obviously has a model, and by compactness, therefore, the whole new theory has a model. But this model also makes up a model of Peano arithmetic itself, of course, and, as constructed, contains an element that is larger than 0, 1, 2, and so on. Thus it is nonstandard.

One can raise the question, then, of how we manage to distinguish between standard and nonstandard models, and especially how we single out the standard ones which, as we are prone to say, are the intended ones that we actually "have in mind".—This is our problem.

Let me briefly explain the terms "standard" and "intended" as used in its formulation. The natural and, as it were, canonical explication that model theorists give of the standard models of arithmetic is the following: The *standard models* are those in which no element has infinitely many predecessors. And all other models of arithmetic—of which there are very many, even very many countable ones—are called *nonstandard*. Normally, however, it is only the standard models that are 'intended', and model theorists also give explications of this term. Wilfrid Hodges, for example, explains it as follows, taking up the historical development in geometry:

> [I]n geometry, axioms were first used for describing a particular structure, not for defining a class of structures. When a theory T is written down in order to describe a particular structure A, we say that A is the **intended model** of T. It often happens—as it did in geometry—that people decide to take an interest in the unintended models too. (Hodges 1997, 32)

So far, so good. But where is our *problem* now? Model theorists do not seem to see any difficulty in singling out the standard models and in explaining what is meant by calling these models "intended" or "not intended", as the case may be! What is the *philosophical* problem that people nevertheless diagnose at this point?

This problem is not easy to express, and I will actually try to show that it turns out to be a pseudo-problem. But in first approximation I would say that it is a sort of dilemma. I call it the *aboutness dilemma* (which is a barbaric term, but it is comfortably short). This dilemma affects almost every mathematical theory—and maybe even empirical theories, as claimed by Putnam, for example[9]—but I will stick to arithmetic alone. Its *first horn* states that if one gives the aboutness of arithmetic a precise formulation that makes it transparent, as is done by means of the model-theoretic notion of interpretation, then we cannot catch it uniquely; we are confronted with a multitude of non-intended interpretations that cannot be avoided. The *second horn* claims that this non-uniqueness *can* be avoided in the meta-language in which model theory is expressed, and typically expressed in a non-formalized way, as just explained for the terms "standard model" and "intended model".[10] But then the aboutness of arithmetic is *not* made precise and transparent; it rather remains in the dark. So, for short: if the aboutness is clear and precise, it is not unique, and if it is unique, it is neither clear nor precise. Both of these possibilities seem unattractive, and the question arises how we might escape from this dilemma without adopting one of them.

It is of no avail in this context to bring the model-theoretic meta-theory into a mathematized form, with the help of a meta-meta-theory, because this threatens to lead into a regress of ever higher meta-theories. Likewise, it doesn't help to take refuge in second or higher order formulations, because the second-order quantifiers are in need of being appropriately understood, and the proponents of second-order logic are ready to admit that their view *presupposes* an intended *standard* interpretation of their quantifiers.[11] The question of how this presupposition is to be substantiated remains totally open when we simply move to second- and higher-order theories. For this reason, I will stick to first-order theories; and this is the common position in discussions about the present problem. Of course, one could say a lot more about this reason, and also about the steps to formalized meta-theories, but this is not my subject in the present paper.

9. See Putnam 1980.

10. It is especially important to have in mind that the characterization of the standard models given before—as those models in which no element has infinitely many predecessors—is formulated in the meta-language.

11. See, e.g., Stewart Shapiro, in Shapiro 1991, 218; Shapiro 1999, 45f. and 58.

To solve our problem, formal devices do not seem to be of much use from the outset. The reason is that formal approaches disregard what we actually think. Don't we simply know what we *mean* by the "standard model"? And don't we know that without making use of any formalizations? Don't we have the firm impression that already now, without further formalizations, we "grasp" what a "standard model" is, in contrast to the nonstandard ones that we dub as "non-intended"? Instead of saying that they *know* what they mean, people also say that they have an *intuition* of what the standard model has to be and that no formalization can capture this intuition. And one can even get the impression that it is precisely the process of formalization itself that is the culprit here; that it actually destroys our understanding of what specifies the standard model.

However, the *problem*, then, is to explain what this "meaning the standard models", this "intuition of them", and what the respective sort of knowledge about them really consists in. As already said, this problem is felt and has been expressed by many people. It is described, for example, in a straightforward way—in a marvellously naïve way, I would say—by the mathematician Yuri Manin on p. 69 of his book *A Course in Mathematical Logic*. There, Manin speaks about the continuum, that is, the domain of real numbers, which of course is affected by precisely the same problem as the domain of natural numbers: our usual first-order theory of the real numbers has non-standard models, in particular it has countable models, and how might they be sorted out? Here is Manin's comment on the situation:

> From the point of view of the topologist or analyst, for whom the continuum is a working reality, the existence of countable models [of topology or analysis] means that formal language has limitations as a means of imitating intuitive reasoning. [...]
>
> For the psychologist or philosopher, perhaps the most interesting aspect of the situation is that any mathematician can understand the viewpoint of another mathematician (without having to agree with it). This means that what mathematician A says, though demonstrably incapable of conveying unambiguous information about the continuum, nevertheless is capable of bringing the brain of mathematician B to the point where it forms an idea of the continuum which adequately represents the idea in A's brain. Then B is still free to reject this idea.

I'm not sure that I totally understand what Manin says here, but I interpret him as follows: From the point of view of the psychologist or the philosopher, the communication between person A and person B, that is, their mutual understanding, is based on the formal linguistic representations realized in their brains; but these representations are nevertheless, as Manin says, "demonstrably incapable of conveying unambiguous information about the continuum". These representations are ambiguous because they do not exclude nonstandard models. But the topologists and analysts then claim that their "intuitive reasoning", as Manin says, is able to remove this ambiguity by going beyond what formal languages convey.

The aim of my quotation of this passage by Manin is not to expose him but to document that our problem is not one that has been concocted by philosophers. It is a problem that also worries professional mathematicians, even such brilliant ones as Manin. And his point of view—that formal languages may be deeply insufficient to capture our intuitions—is very suggestive indeed and has been adopted by many people.[12] But nobody gives a satisfying explication of the notion of "intuition" here. The supposed faculty of intuition remains in the dark, and our problem—the aboutness dilemma—still persists.

At the same time, it seems to me that the term "intuition" should not be discarded.[13] With this notion, Manin and others want to capture

12. For example, it seems to be Paul Bernays' position as he expressed it in a discussion with Thoralf Skolem in 1938 (as it is described in Bellotti 2006, 182f., and recounted in Mühlhölzer forthcoming). Bernays thinks that the purely formal axiomatic method is insufficient and that what we need are notions of "number" and "set" in their "intuitive sense". "Intuitive" is the catch-word here. Allegedly, we possess an intuition of the standard models and this intuition should sort out the nonstandard ones. In accord with this point of view, Bernays emphasizes that the set theory presupposed in model theory should *not* be considered in its formalized version but as what one calls "naïve" set theory. With the help of sufficient restrictions of its use such a theory can very well avoid the known set theoretic antinomies (as is done, for example, in Hausdorff 1914 or in Halmos 1960). Bernays emphasizes that "we need such an intuitive handling of set theory" ("anschauliche Handhabung der Mengenlehre") when "set theory is not only supposed to represent a certain structure, but also to deliver the method of our thinking about structures" (my translation; see Bernays 1971, 200: "Daß wir einer solchen anschaulichen Handhabung der Mengenlehre bedürfen, ist verständlich, sofern die Mengenlehre nicht nur eine gewisse Struktur repräsentieren, sondern auch die Methode unseres Denkens über Strukturen liefern soll"). Interestingly, Bernays not only talks about the "intuitive handling of set theory", but also of "the intuitive handling of set theoretic ideas" ("die anschauliche Handhabung der mengentheoretischen Vorstellungen"; my translation; see Bernays 1971, 213.) It is not merely the symbols of the theory which are handled, but also the ideas expressed by these symbols.

13. If I see it correctly, this is done in the course of Putnam's reflections in Putnam 1980. Michael Dummett tries to save the notion of intuition with regard to models, but he interprets it in a way that I do not find very helpful. This is his proposal: "An intuitive model is a half-formed conception of how to determine truth-conditions for a given class of sentences. […] It is merely an idea in the embryonic stage, before we have succeeded in the laborious task of bringing it to

what is not represented in formalizations, and we might try to interpret it thus by looking at our actual mathematical *practice*, at our actual *use* of relevant terms. A manoeuvre of this sort, albeit in a context different from the context of my present problem, has been nicely enunciated by the economist Piero Sraffa in notes from March 1934 that are connected to discussions with Ludwig Wittgenstein. Sraffa writes:

> The error is to regard intuition as a provisional substitute for science: 'when you will produce a satisfactory science, I shall give up intuitions'.—Now the two things cannot be set against one another they are on entirely different planes. Intuitions are a way of acting, science one of knowing.[14]

Let us call this *Sraffa's idea* of explicating "intuition" via "acting", that is: via "use". Under this perspective, Manin's talk about "intuition" can get a really substantial and pleasantly down-to-earth interpretation: it can be understood as an allusion to our familiar use of the pertinent terms. This familiar use does not involve strict formalizations; in fact, it has developed over centuries during which formalization in our present sense—formalization that comes to its summit in model theory—was simply not thought of or even not thinkable. And to this not-regimented use we should now turn in order to approach our problem. To my mind, it is the key to a solution—or a dissolution—of the problem.

The use that should interest us has mainly to do with the deep difference between our *referring* to numbers in our actual mathematical practice, on the one hand, and our *interpretation* of a formal language, on the other hand, as is done in model theory. The Wittgensteinian solution of the problem that I will present consists in pointing to precisely this difference. In the rest of this paper, I will try to explain it in detail, but first let me roughly state the solution, or better: dissolution, in a few sentences. It consists in the following train of thoughts:

birth in a fully explicit form." (Dummett 1967, 214) I do not think that this is what Bernays and Manin had in mind. They used the term "intuition" in order to catch everything essential that is factored out in model-theoretic formalizations, without hoping that future formalizations might be of any avail.

14. In McGuinness 2008, 229.—A trace of this idea may be seen in RFM IV §32, where Wittgenstein writes: "I am asking: what is the characteristic demeanour of human beings who 'have insight into' something 'immediately', whatever the practical result of this insight is? | What interests me is not having immediate insight into a truth, but the phenomenon of immediate insight. Not indeed as a special mental phenomenon, but as one of human action." Instead of "immediate insight", Wittgenstein here also might have said "intuition". But the subject "intuition" is actually a polymorphic one in Wittgenstein's work, and I cannot pursue it further.

Our problem is muddled because in its usual formulations the distinction between "reference" and "interpretation" is blurred. The quite common talk about an 'intuition' that one has *about the standard models*, an intuition that goes beyond what model theory can afford, is confused because it belongs to the domain of used signs within which the notion of "reference" has its place, and it does not involve the notions of "interpretation" and "model" as they occur in model theory, which concern only signs that are considered as purely mathematical entities and are not used. Our alleged "intuition" of standard models belongs to the realm of used signs, and it concerns reference and not interpretation. That is, it has nothing specific to do with models in the model theoretic sense.[15] Our knowledge of standard models and of the difference between them and the nonstandard ones is a purely mathematical one and does not raise any specific problems that go beyond problems concerning mathematical knowledge in general. To think otherwise is simply a delusion, and this delusion is the basis of the alleged dilemma concerning aboutness. The aboutness of its first horn is model-theoretic interpretation, and we should not be worried about its non-uniqueness; and the aboutness of its second horn is something different: it is reference, which can be sufficiently clarified (as I will show in the following). In other words: our problem is a pseudo-problem because the alleged tension between these two horns is a delusion.

This is the way I would dissolve the aboutness dilemma—and what I find amazing now is the ubiquity of the conflation of "reference" and "interpretation" that pervades the literature concerning this dilemma. Let me briefly come back to the already quoted passage of Manin, which is a perfect example for this. When claiming that mathematicians A and B, thinking and talking about the continuum, are "demonstrably incapable of conveying unambiguous information about the continuum", Manin is the victim of the confusion mentioned. What is *demonstrably* ambiguous is our talking about the *model* of our theory of the continuum, because we have mathematical proofs that there exist nonstandard models; but in this sense mathematicians are not incapable of anything. They are very well capable of distinguishing between standard and nonstandard models *in* model theory. And this does not involve any *intuition*. When Manin brings 'intuition' into play, in the case of (as he says) "the topologist or analyst, for whom the continuum is a working reality", this is not done

15. This contrast between used and non-used signs is elaborated also in Mühlhölzer 2012 and Mühlhölzer forthcoming.

from the model theoretic point of view but from the standpoint of the practicing mathematician whose signs are used to *refer* to his familiar points and numbers, in contrast to the non-used signs and their *interpretations* considered in model theory. Manin does not sufficiently distinguish between these two domains, and this is the main defect of his reflections.

The same mistake can be diagnosed in Putnam 1979, when on pp. 118f., under the heading "Wittgenstein on 'following a rule'", Wittgenstein's rule-following considerations are connected with the mathematical fact that one "may have divergent interpretations of the whole theory, as the Skolem-Löwenheim theorem shows" (Putnam 1979, 119). But 'following a rule' in Wittgenstein's sense essentially concerns our use of the signs involved, whereas the so-called 'signs' that are interpreted in model theory, with the Löwenheim-Skolem indeterminacy as a mathematical consequence, are not considered as used at all. They are purely mathematical entities, light years away from what Wittgenstein had in mind when talking about rules.

3. *The partial and the full use thesis concerning reference*

The distinction I have in mind, between "reference" and "interpretation", must now be clarified. Let us begin with "reference". As I understand this term here, it is a generic notion that incorporates *naming* in the case of singular terms (as naming the number 3 with the respective numeral), *to apply to* in the case of general terms (to apply to the number 7, say, in case of the general term "prime number"), and *aboutness* in the case of statements (as when one says, for example, that »7 + 5 = 12« is a statement about numbers). And one should speak of 'reference' only in the case of *used* expressions. Reference essentially depends on use and a philosophical investigation of reference must be an investigation of use. In the literature one comes across the so-called *use thesis* concerning reference. It says that in the determination of the reference of our terms the use we make of the terms must play an essential part, and this thesis has been widely adopted. If I see it correctly, the term *use thesis* was originally introduced by Stewart Shapiro, not, however, as a thesis about reference, and also not about meaning in general, but about *understanding* a language. He explains it thus: "The claim is that understanding should not be ineffable. One understands the concepts embodied in a language to the extent that she knows how to use the language correctly. Call this the *use thesis*." (Shapiro 1990, 252; similarly in Shapiro 1991, 211–214, and 1997, 204ff.) Shapiro seems to connect this thesis with "reference", but the way he does it is not clear to me. It is Hilary Putnam who explicitly connects the use thesis with reference. This is what he says: "On any view, the understanding of the language must determine the reference of the terms, or, rather, must determine the reference given the context of use." (Putnam 1980, 24) In what follows the term "use thesis" is exclusively meant as such a thesis about reference.

Normally, the determination of reference by use is regarded as a merely partial one—factors other than use may play a role as well –, and I call a thesis of this sort the *partial use thesis*.[16] Often one associates with it the idea that our use may *help* to constitute the necessary *link* between ourselves and the objects referred to, such that it might become possible to *explain* how we are *able* to refer. So, Harold Hodes says:

> [W]hatever the links between ourselves, our practices and abilities, and objects like Jupiter Exxon, tyrannosaurus rex, and positrons, by virtue of which we and our words refer to them, these links are going to have to be rather different from any such links between ourselves, our practices, our abilities and the number 1. Numbers are so pure, so unsustained by the cement of the universe, that reference to them and their ilk seems quite *sui generis*. (Hodes 1984, 127)

The idea here seems to be that the reference of our words is based on such 'links', or consists in them, and that the task of philosophy is now to ferret them out. A profound task, as it seems! Tim Button and Peter Smith describe it as follows:

> [S]ince we are little more than shaved apes, wearing shoes, how can we possibly manage to refer to abstract objects (numbers or otherwise)?
>
> There is much to be said on *this* question, of course, but not here. (Button/Smith 2012, 120)

One must say, however, that with such views the precise nature of our 'managing to refer' remains in the dark. People think that this nature may be made transparent in the future, but at present it is still in the dark. Accordingly, our *aboutness dilemma* persists: the aboutness of arithmetic, when understood via reference according to this conception, is still in the dark, and the formal clarification in model theory destroys its unique-

16. This is the way in which Charles Travis comments upon PU §432: "Use, the suggestion is, is a factor on which the semantics of an item depends." (Travis 1989, 17) As I will say in a moment, I do not think that this is a correct understanding of Wittgenstein's thinking.

ness.—What should we do in this situation? Simply wait until we have a respectable philosophical theory about 'reference', a theory that uncovers the link between language and objects with which a philosophical explanation of our ability to refer can be given?

I do not believe that such a theory can be developed, and instead I would propose to make the aboutness of arithmetic transparent by adopting a stricter version of the use thesis, a version which I call the *full use thesis*. To my mind, the use thesis is ambiguous in an essential way—it can be understood as the *partial* or as the *full* use thesis –, and our dilemma should be resolved by concentrating on the aboutness of arithmetic in a sense of "reference" conforming to the full use thesis. It says that *there is nothing more to reference than what can be seen in the use of our terms*. This use doesn't help us to disclose a hidden essential 'link' between ourselves and the objects referred to, and it cannot 'explain' how we are 'able to refer'. According to the Wittgensteinian perspective, when in the second horn of our aboutness dilemma one says that the aboutness of arithmetic 'remains in the dark' and thereby means the reference of the arithmetical terms, one gives 'reference' a wrong emphasis. From the outset, there is nothing in the dark with it, but everything is in plain view. Our talking about 'reference' merely brings out certain aspects of our use that we deem to be important in certain circumstances.

It is such a view of 'reference' that I see expressed by Wittgenstein in the already mentioned and partly quoted §10 of the *Philosophical Investigations*. Wittgenstein, of course, would not call this view a "thesis", and he does not use a general and somewhat artificial term like "reference" which seems to hint at a philosophical theory about reference that has not yet been found. Instead, as already said, in the original German version of §10 Wittgenstein uses the word "bezeichnen", which is a very ordinary word in German. As already said as well, in the present context I will stick to the word "refer" in order to connect Wittgenstein's view with the way actual discussions in the philosophy of mathematics are conducted. And for this reason I will also keep to using the word "thesis". What will remain essential, however, is the fact that with "reference" and its cognates Wittgenstein doesn't associate deep *links* that connect words with objects and that go beyond, or constitute a sort of basis, of what our use of the words involves. This was different in the *Tractatus*. There he had considered words like these—and also words like "meaning" and "understanding"—as 'metalogical' ones, as expressing concepts which are essentially presupposed by logic and therefore have a particularly profound status (see Hacker 1996, 685 and 688). In his later philosophy, however, he consequently deprives them of that.[17]

The view that use is primary over reference throughout—what I call the "full use thesis"—does not lose sight of *objects* and the way language is *about* objects. Wittgenstein doesn't reject calling numbers "objects", he only warns of giving this manner of speaking philosophical weight. He says this as follows in MS 116, 254f.:

> One doesn't change the grammar of the word "the number 3" by calling it [= the respective number] 'an object'. We thereby merely consent to say "the object three" instead of "the three". Again and again one confuses propositions that propose a new picture, a new name, with propositions about the nature of an object. If I introduce a new figurative sort of expression: I must transfer, then, the grammar of the old expression to the new one.

When talking about words' 'being about objects' we may be inclined to think of specific 'connections between words and objects', but these so-called 'connections' constantly change from case to case and from context to context, and each time they have to be seen in the manner we *use* our words. And when using the word "reference" itself, or one of its cognates, we highlight aspects of these so-called connections. Also the use of the word "reference" and its cognates can be very different for different cases and contexts, and to think that one may find something uniform on the basis of all this—the 'universal mechanism of reference', so to speak, which ultimately picks out the objects—is a philosophical dream that will not come true. According to the Wittgensteinian perspective, it is a philosophical fantasy.

Incidentally, Wittgenstein's view appears especially appropriate in the case of mathematics where from the outset we are reluctant to think that the objects we refer to with our symbols are given in advance and that the use we make of the symbols should then be explained according to a prefabricated connection between the symbols and the objects. (See Mühlhölzer 2012, where this point of view is explained in more detail and where consequences are drawn with regard to metamathematics.) The

17. See the motto of this paper, from MS 116, 3. On p. 16 of this manuscript Wittgenstein explicitly says: "For a long time I was tempted to believe that »understanding« is a *metalogical* word". On p. 57f. he also mentions the word "rule" as one that should be considered as a word like all others, and he adds the general remark: "[I]n philosophy things are going on in quite a gemuetlich/leisurely manner; we do not construct a (unique) monumental structure, *the* proper *language*, but want to remove misunderstandings".

full use thesis, the view that use is primary over reference, is particularly attractive in the case of mathematics!

So, instead of searching for a hidden basis ultimately constituting reference, we should look at the function of our actual talking about 'reference'. The function mentioned by Wittgenstein in §10 of the *Philosophical Investigations* is only: to remove certain misunderstandings; but, of course, other functions can be easily stated. In the context of the aboutness dilemma, however, which is our present problem, we may diagnose something like a misunderstanding as well, namely: to confuse our *referring* to numbers when using number-words with the *interpretations* of these words when they are considered as signs of the object language of model theory where they are no longer used. Precisely this is my topic here. It has not been considered by Wittgenstein himself, but it is tailor-made for a Wittgensteinian approach.

4. *Interpretation versus reference*

In our usual practice of using number words we 'refer to our familiar numbers', as we can say. That is, our normal way of using the word "reference" and its cognates allows us to talk in this way, and there is no reason to turn it down.[18] In model theory, however, "reference" is now thoroughly *replaced* by "interpretation", "interpretation" in the sense of interpreting *formal languages* by means of certain structures. In what follows, I will use the term "interpretation" only in this sense. And the point I want to make is that interpretation in this sense must not be identified with reference; that there is, in fact, a categorial difference between these two notions. The blurring of this difference is the core of our problem.

Of course, in the *meta*language in which model theory is conducted—that is, which is *used* in model theory—we *refer* to mathematical entities, and we can even think of referring to our familiar numbers in this metalanguage. Normally, however, the metalanguage actually presupposed is the language of set theory, and our familiar numbers are represented by their set theoretic proxies. Furthermore, the models referred to in model

18. In particular, our usual mathematical practice is not a foundationalist one. In a foundationalist context mathematicians may wish to refer with the numeral "3" to a set theoretic proxy of the familiar number, like the corresponding von Neumann number or the Zermelo number, but one should be aware that this is only a sort of simulation in set theory, which need not—and which should not—be identified with the number referred to in our mathematical practice.

theory are often, and quite characteristically, constructed out of the 'signs' of the object languages considered.[19] So our normal ways of referring to numbers aren't present in model theory as it is actually done. In its actual practice we tend to refer to rather artificial entities which, however, are the appropriate ones for the theoretical purposes at hand.

The decisive point with respect to "interpretation" in the model theoretic sense is that the interpreted 'language'—the object language of model theory—is not a language considered as used. It is a purely mathematical object, and mathematical objects are not used: only the signs with which we refer to these objects are used. Precisely this use is totally absent in the case of an interpretation. Of course, as just said, the metalanguage is used—but not the object language. Such a situation can be described with a felicitous metaphor that is sometimes employed by Wittgenstein[20]: in the case of an interpretation, the object language that is interpreted is *dead* like any other mathematical object, and only the meta-language which is about the object language has (or may have) life. In order to refer, the signs of a language must have life; in order to be interpreted, they must be dead.[21]

Seen from this point of view, "interpretation" appears categorially different from "reference". And this the more so if we consider two further aspects of "interpretation" that make it different from "reference". First, there is the issue of *existence*: the existence of an interpretation is normally settled in a purely mathematical way, while the question of whether terms refer seems to have quite another, genuinely philosophical status. Second, interpretation in the model theoretic sense, in contrast to reference, functions as an essential mathematical tool, as part of techniques or methods that aim at solving specific mathematical problems. This is already true of Hilbert's use of interpretations in order to get metamathematical results, but it has been much expanded in the development of mathematics after

19. Hodges has a special name for models of this sort: he calls them "canonical models" (see Hodges 1997, 19).

20. For example in BB, 3ff., and PI §§432 and 454.

21. See on this point also Mühlhölzer forthcoming, to which the present paper is a sort of sequel, albeit a sequel that can be understood in itself, I hope.—To my mind, one shouldn't be bothered by this metaphorical manner of speech. Someone might object: "Why not talk about 'meaning' instead of 'life'?" But would this really be an improvement? It suggests a more respectable dealing with language, or even a *theory* of meaning, but this might prove to be an illusion in the end. And even if such a theory might be realized in the future: should we wait until then? And will it be helpful for our attempts at treating philosophical problems? From a Wittgensteinian standpoint, we should not wait but try to tackle philosophical problems now; try to show, for example, that his perspective is helpful for overcoming the aboutness dilemma.

Hilbert, where we can also find results belonging to mainstream mathematics that are proved model theoretically and that are now the model theorists' pride and joy. Nothing comparable can be said with respect to the notion of reference.

To say it pointedly: these so-called 'signs' of an object language of model theory do not *refer to* anything; to use the term "reference", or cognate terms like "signify", etc., is a category mistake if we mean by "reference", etc., what we normally mean by these words in the context of reflections about language. The 'signs' of model theoretic object languages do not 'refer', just as numbers do not refer.

One might object that I now have exaggerated the distance between reference and interpretation, and to a certain extent this is true. One might think, for example, that there exists a wider notion of "interpretation", a notion that in a sense comprises model theoretic "interpretation" *and* "reference", which I missed so far.—But the sense of this wider notion should then be clarified. From the perspective that I have described up to now, it seems that this wider notion must involve a *transition* from the dead 'interpretations' of model theory to the life that must be seen in what is called "reference". And the problem here is to understand this transition. Might it be the case, for example, that interpretation is some sort of formal core of reference, or "a mere shell of the reference relation", as Shapiro says (Shapiro 1997, 139)? And that reference, as the richer thing, might accomplish what this core or shell alone cannot? For example, that it might accomplish the selection of the *intended* interpretations out of all the interpretations that constitute models of a given formalized theory? But I do not really understand this philosophical picture. Despite the amount of literature that has been produced in its sphere it remains desperately in need of clarification.

Our puzzle about the singling out of standard models stems from the fact that the *connection* between the domain of live signs, belonging to our mathematical practice, and the domain of dead signs, making up an object language of model theory, remains unclear. So, what should be clarified is this connection, which has been ignored up to now. What does it precisely consist in?

The relevant connection, I think, consists in the way mathematicians *transform* aspects of their practice into formulae of a formal language investigated in model theory. To take an elementary example: our practice of dividing a natural number a by a natural number b and thereby getting a result with no remainder is now expressed by the formula $\exists x(b \times x = a)$.

And so on. In this way aspects of our live practice are, so to speak, petrified in dead formulae.

But it is precisely this characteristic sort of transformation of live into dead formulae which gives rise to nonstandard models and which lets us find out many of their properties.[22] It does not reduce the gap between the language of our live mathematical practice and the object languages of model theory which underlies our philosophical problem. It only shows what the gap precisely consists in.—How deep is this gap? Isn't it merely produced by an *idealization* of our ordinary practice, by one of those standard idealizations that are the staff of life in any science?—No, we should sharply distinguish at this point between idealization and *petrification*, petrification in the sense of making live signs dead. Our normal processes of idealization consist in disregarding certain untidy aspects of real situations that we deem inessential in the context at hand, as when we leave aside small but inevitable occurrences of friction in physics. When petrifying signs, on the other hand, we disregard *what is essential* to them *as signs*, namely their being used, and seen in that way, petrification appears to be almost the opposite of idealization. The phenomenon of petrification is clearly illustrated in the usual proofs of the *completeness* and the *compactness theorem* which guarantee the existence of models of a formalized first-order theory under comparably weak conditions. In these proofs the so-called 'signs' of the theory are treated as mathematical entities—as dead mathematical entities—out of which the models are constructed in a quite straightforward way, a way that only involves the addition of new 'signs' of this dead sort in order to fulfil certain structural needs.[23] What is operative here is not idealization but petrification.

But shouldn't one say that both, reference and interpretation, are specific *relations* between signs and objects?—Of course, one can say that, but it is a totally superficial way of speaking, because these 'relations' are so different. Interpretation is a purely mathematical relation, a mathemati-

22. According to Boolos et al. this transformation is the basis of what they call "the main technique for obtaining information about the 'appearance' of [a nonstandard model] $\mathcal{M}$". They describe this technique as follows: "observe that the natural numbers have a certain property, conclude that a certain sentence of [the formal language L] is true in [the standard model] $\mathcal{N}$ [where "true in a model" is a technical term of model theory], infer that it must also be true in $\mathcal{M}$ (since the same sentences of L are true in $\mathcal{M}$ as in $\mathcal{N}$), and decipher the sentence 'over' $\mathcal{M}$" (Boolos et al. 2007, 202). The transformation in the converse direction—from the dead to the live signs—is sometimes called "reading"; see Ebbinghaus et al. 1994, 30 (the live-to-dead direction is there simply called "formalization": see 44–52).

23. See, for example, Ebbinghaus et al. 1994, 75–90; Hodges 1997, 17–42 and 124ff.

cal function that in the standard set theoretic framework is considered as a pure set, with so-called "signs" that do not deserve this name. Seen in that way, interpretation shows a great uniformity. This is totally different, however, for reference, which is fundamentally tied to use and to the non-uniformity of use. The *partial* use thesis doesn't make that sufficiently clear! This non-uniformity is taken account of in the *full* use thesis, however, and it concerns a fundamental Wittgensteinian point with regard to reference. An important aspect of the Wittgensteinian perspective is that there is *no essence* underlying all cases of reference. The function of the word "reference" and its cognates is multifarious, contrary to what we find with respect to "interpretation".[24]

References

Baker, Gordon & Hacker, Peter 2009: *Wittgenstein: Rules, Grammar and Necessity*. Vol. 2 of *An Analytical Commentary on the* Philosophical Investigations. 2nd edition by Peter Hacker. Oxford: Blackwell.

Bellotti, Luca 2006: "Skolem, the Skolem 'Paradox' and Informal Mathematics". *Theoria* 72, 177–220.

Bernays, Paul 1971: "Zum Symposium über die Grundlagen der Mathematik". In: Paul Bernays (1976), *Abhandlungen zur Philosophie der Mathematik*. Darmstadt: Wissenschaftliche Buchgesellschaft,189–213.

Boolos, George, Burgess, John & Jeffrey, Richard 2007: *Computability and Logic*. 5th edition. Cambridge/New York: Cambridge University Press.

Bourbaki, Nicolas 1942: Élements de mathématique, fascicule iv, livre II. ALGÈBRE, chapitre 1.Structures algébriques. Paris: Hermann.

Button, Tim & Smith, Peter 2012: "The Philosophical Significance of Tennenbaum's Theorem". *Philosophia Mathematica* (III) 20, 114–120.

Diamond, Cora 1996: "Wittgenstein, Mathematics, and Ethics: Resisting the Attraction of Realism". In: Hans Sluga & David Stern (eds.), *The Cambridge Companion to Wittgenstein*. Cambridge: Cambridge University Press, 226–260.

Dummett, Michael 1963: "The Philosophical Significance of Gödel's Theorem". *Ratio* 5, 140-155; reprinted in Dummett 1978, 186–201.

24. I'm grateful to Juliet Floyd for suggestions in the run-up to this paper and to Wilfried Keller, Dolf Rami and the participants of the Conference *Perspectives on Wittgenstein's Philosophy of Mathematics* (University of Zürich, August 2012) for valuable remarks on preliminary versions of it.

Dummett, Michael 1967: "Platonism". First published in Dummett 1978, 202–214.
— 1978: *Truth and Other Enigmas*. London: Duckworth.
Ebbinghaus, Heinz-Dieter, Flum, Jörg & Thomas, Wolfgang 1994: *Mathematical Logic*, 2nd edition. New York: Springer.
Hacker, Peter 1996: *Wittgenstein: Mind and Will*. Vol. 4 of *An Analytical Commentary on the* Philosophical Investigations. Oxford: Blackwell.
Halbach, Volker & Horsten, Leon 2005: "Computational Structuralism". *Philosophia Mathematica* (III) 13, 174–186.
Halmos, Paul 1960: *Naive Set Theory*. New York: Springer.
Hausdorff, Felix 1914: *Grundzüge der Mengenlehre*. Leipzig: Von Veit.
Hodes, Harold 1984: "Logicism and the Ontological Commitments of Arithmetic". *The Journal of Philosophy* 81, 123–149.
Hodges, Wilfrid 1997: *A Shorter Model Theory*. Cambridge: Cambridge University Press.
Manin, Yu. I. 1977: *A Course in Mathematical Logic*. New York: Springer.
McGuinness, Brian (ed.) 2008: *Wittgenstein in Cambridge: Letters and Documents 1911-1951*. Oxford: Blackwell.
Mühlhölzer, Felix 2012: "Wittgenstein and Metamathematics". In: Pirmin Stekeler-Weithofer (ed.), *Wittgenstein: Zu Philosophie und Wissenschaft*. Hamburg: Felix Meiner Verlag, 103–128.
— forthcoming: "On Life and Dead Signs in Mathematics". In: Godehard Link (ed.), *Formalism and Beyond. On the Nature of Mathematical Discourse*. Frankfurt a. M.: Ontos Verlag.
Putnam, Hilary 1979: "Analyticity and Apriority: Beyond Wittgenstein and Quine", as reprinted in Putnam 1983, 115–138.
— 1980: "Models and Reality", as reprinted in Putnam 1983, 1–22.
— 1983: *Realism and Reason. Philosophical Papers, Vol. 3*. New York: Cambridge University Press.
Shapiro, Stewart 1990: "Second-Order Logic, Foundations, and Rules". *The Journal of Philosophy* 87, 234–261.
— 1991: *Foundations without Foundationalism: A Case for Second-order Logic*. New York: Oxford University Press.
— 1997: *Philosophy of Mathematics: Structure and Ontology*. New York: Oxford University Press.
Travis, Charles 1989: *The Uses of Sense. Wittgenstein's Philosophy of Language*. New York: Oxford University Press.

Grazer Philosophische Studien
89 (2014), 61–78.

WITTGENSTEIN ON EQUINUMEROSITY AND SURVEYABILITY

Mathieu MARION
Université du Québec à Montréal
&
Mitsuhiro OKADA
Keio University

Summary

This paper aims to connect two of Wittgenstein's arguments against Logicism. The 'modality argument' is directed at the Frege/Russell-definition of numbers in terms of one-one correlations. According to this argument, it is only when the Fs and Gs are few in number that one can know that they can be one-one correlated without knowing their numbers. Wittgenstein's 'surveyability argument' purports to show that only a limited portion of arithmetic can actually be proven within *Principia Mathematica*. For proof-constructions within this system quickly become unsurveyable and thereby loose their cogency. As we shall argue, the role of visualisation in proofs plays a fundamental role in both arguments.

1. *Introduction*

In this paper, we wish to draw attention to a close link between two important arguments in Wittgenstein's philosophy of mathematics, namely his argument against the use of the notion of one-one correlation in the Frege-Russell definition of numbers, which we call here the 'modality argument', and his notorious 'surveyability argument'. The fact that these two arguments are closely related points to a nexus in his philosophy of mathematics, a point where it would stand or fall. The former argument has received less attention in the secondary literature, but both have at any rate failed to convince. Alas, this is not the place to mount a defence, should we wish to, because it is important first to understand correctly the nature of those arguments. We limit ourselves here to this last task only—but this task should be seen, however, as part of a larger investigation of the core claims of Wittgenstein's philosophy of mathematics. We shall thus first propose in some detail an interpretation of the modality argument and then briefly show how it is related to the surveyability argument.

2. *The modality argument*

It is useful, in order to understand the point of Wittgenstein's argument against one-one correlation, to recall some details of the Frege-Russell definition of numbers. In §63 of *Grundlagen der Arithmetik*, Frege introduced a cardinality operator, 'the number of Fs', which is nowadays written:

$$\mathrm{N}x : \mathrm{F}x.$$

He introduced this operator with a contextual definition, which is known as 'Hume's principle', according to which 'the number of Fs is equal to the number of Gs if and only if they are in a one-one correlation' (here: $\mathrm{F} \approx \mathrm{G}$):

$$\mathrm{F} \approx \mathrm{G} \leftrightarrow \mathrm{N}x : \mathrm{F}x = \mathrm{N}x : \mathrm{G}x$$

The idea behind this definition is that it provides an all important criterion of identity, i.e., a criterion for our being able to recognize again the same number. The key here is thus Frege's definition of 'equinumerosity' (§§71–72), which reads like this: F and G are 'equinumerous' just in case there is a relation R such that every object belonging to F—Frege would say 'falling under F'—has the relation R to a unique object belonging to G and every object falling under G is such that there is a unique object belonging to F which also has the relation R to it. To get the definition going, we need the notion of 'unique existence':

$$\exists! x\, \mathrm{H}x =_{\mathrm{def}} \exists x\, (\mathrm{H}x \;\&\; \forall y\, (\mathrm{H}y \rightarrow y = x))$$

So the definition reads formally as:

$$\mathrm{F} \approx \mathrm{G} =_{\mathrm{def}} \exists \mathrm{R}\, ((\forall x\, (\mathrm{F}x \rightarrow \exists! y\, (\mathrm{G}y \;\&\; \mathrm{R}xy)) \;\&\; (\forall x\, (\mathrm{G}x \rightarrow \exists! y\, (\mathrm{F}y \;\&\; \mathrm{R}xy)))$$ [1]

1. We omit details that are of no importance here, e.g., the fact that one can show that '≈' is reflexive, symmetric and transitive, etc.

With this notion Frege can then define, in §73, the number of Fs in terms of 'classes of classes':

$$Nx : Fx =_{\text{def}} \{G: G \approx F\}$$

And from there he can go on defining natural numbers. For example, a number such as 2 is defined in terms of the class of all classes that are in one-one correlation with a given pair. In *Principia Mathematica*, Russell and Whitehead proceed in a similar fashion to obtain the same definition, albeit in the rather complicated syntax of their type theory. Hence the name 'Frege-Russell definition'—it is a key to their 'logicism'.

Hume's principle is a biconditional, but Frege provides an argument that might properly be called 'philosophical' to the effect that the direction that really counts is from left to right, i.e., from the fact that there is a one-one correlation '$F \approx G$' to the sameness of number, or '$Nx : Fx = Nx : Gx$'. Frege's argument at §§64–68 involves, however, showing the priority of 'The line *a* is parallel to *b*' over 'The direction of *a* is the same as the direction of *b*', and this gets him into some further difficulties into which we need not get into. The reason for his having to provide an argument here is in the end rather simple: in order for Hume's principle to serve in a convincing manner for the definition of natural numbers, one must, for fear of circularity, use some other notion that does not involve numbers; one-one correlation, he argues, is just this *prior* notion. Indeed, one can correlate a bunch of cups and saucers to see that they are equal in number, without knowing what that number is. To find out what that number is, one would correlate them with natural numbers, i.e., count them.

Wittgenstein discussed the Frege-Russell definition on numerous occasions in his writings and lectures, from the 'middle period' up to and including the 1939 lectures on the foundations of mathematics.[2] Among his numerous remarks, the modality argument plays a central role.[3] One early occurrence of it is in Wittgenstein's conversations with Schlick and Waismann (January 1931):

2. See LFM, 157f. For this reason it would be wrong to dismiss his remarks as merely pertaining to the apparently discredited 'middle period', a typical but exegetically unwarranted move.

3. For the modality argument itself, see Wittgenstein 2003, 373f.; WVC, 164f.; PR, §118; BT, 415; PG, 355f.; AWL, 148f., 158, 161ff. One should note that Wittgenstein hardly ever refers to Frege, but discusses at length the specifics of Russell's own version. There is no need to get into this, however, within the context of this paper.

In Cambridge[4] I explained the matter to my audience in this way: Imagine I have a dozen cups. Now I wish to tell you that I have got just as many spoons. How can I do it?

If I had wanted to say that I allotted one spoon to each cup, I would not have expressed what I meant by saying that I have just as many spoons as cups. Thus it will be better for me to say, I can allot the spoons to the cups. What does the word "Can" mean here? If I meant it in the physical sense, that is to say, if I mean that I have the physical strength to distribute the spoons among the cups—then you would tell me, We already knew that you were able to do that. What I mean is obviously this: I can allot the spoons to the cups because there is the right number of spoons. But to explain this I must presuppose the concept of number. It is not the case that a correlation defines number; rather, number makes a correlation possible. This is why you cannot explain number by means of correlation (equinumerosity). You must not explain number by means of correlation; you can explain it by means of possible correlation, and this precisely presupposes number.

You cannot rest the concept of number upon correlation. [...] When Frege and Russell attempt to define number through correlation, the following has to be said:

A correlation only obtains if it has been *produced*. Frege thought that if two sets have equally many members, then there is already a correlation too... Nothing of the sort! A correlation is there only when I actually correlate the sets, i.e. as soon as I specify a definitive relation. But if in this whole chain of reasoning the *possibility* of correlation is meant, then it presupposes precisely the concept of number. Thus there is nothing at all to be gained by the attempt to base number on correlation.[5]

We have to keep in mind when interpreting Wittgenstein that the Frege-Russell definition of number is in terms of logic, that is in terms of 'classes' and 'objects' that belong to them. A one-one correlation is thus meant to be a pairing of these objects.

With this point kept in mind, the modality argument is as follows: it is not the case that *there always is* a one-one correlation, as defined above, between the objects belonging to any two classes with the same number (of objects belonging to them). Of course, there *could* be such a correlation between any two classes with the same number of objects belonging to them. So one may claim that any such a correlation not yet

4. The minutes of the Trinity Mathematical Society, reproduced at Wittgenstein 2003, 373, show that Wittgenstein discussed this very topic during their meeting on May 28, 1930.

5. WVC, 164f.

established *can* always be established. One might counter this last move, however, by pointing out with Louis Goodstein that this 'can' is only a "logical possibility",[6] and that this possibility looks more like the *consequence of* the fact that the two classes have the same number of objects belonging to them, than a *condition for* them to have the same number of objects belonging to them. However, if one *already knows* that the two classes have the same number of objects belonging to them, then one surely *knows* that a one-one correlation *can* be established. So, the argument goes, Frege's philosophical claim for the priority of one-one correlation does not hold, and the definition is in danger of simply being circular. One could, however, point out that circular definitions abound in mathematics and that they are not necessarily vicious, so that the claim that the Frege-Russell procedure is in the end circular cannot be held against it without further justification. But, as we said, we do not wish to get side-tracked at this stage into issues pertaining to the evaluation of the argument.

The modality argument also occurs in the writings of Friedrich Waismann[7] and Louis Goodstein,[8] but in both cases one can argue that the idea originates in Wittgenstein.[9] As Michael Dummett once pointed out, "very few objections [...] have ever been raised" (Dummett 1991, 148) against the Frege-Russell definition, so the modality argument is for that reason of intrinsic importance, even if it is ultimately deemed a failure. But, apart from a short discussion of Waismann's version by Dummett,[10] it has attracted surprisingly little attention.[11] This fact might be explained

---

6. See Goodstein 1951, 19. This is also strongly implied in BT, 415; PG, 356.

7. See Waismann 1951, 108f.; and Waismann 1982, 45f.

8. See Goodstein 1951, 19.

9. In an 'Epilogue' to *Introduction to Mathematical Thinking*, Waismann identified a manuscript by Wittgenstein (possibly the manuscript now published as *Philosophical Remarks*) as the source for his argument (Waismann 1951, 245); we just saw that he knew the argument from a conversation with Schlick and Wittgenstein in 1931 that he recorded himself in Gabelsberger shorthand. As for Goodstein, he does not give any indication, but the fact that he had been a student of Wittgenstein in the early 1930s, who was largely inspired by him in his own work in mathematical logic, leads us to believe that Wittgenstein is again the source here.

10. See Dummett 1991, 148f.

11. While one of us was probably the first to attract attention to it (Marion 1998, 77–83) there is to our knowledge only a short abstract by Daniel Isaacson (Isaacson 1993), an interesting pair of papers by Boudewijn de Bruin (de Bruin 1999) and (de Bruin 2008), and a short discussion by Gregory Landini in *Wittgenstein's Apprenticeship with Russell* (Landini 2007, 168ff.). (Although Isaacson 1993 was in fact published earlier, it was prompted by Marion 1991, 81–88, which eventually found its way, in a revised form, in Marion 1998, 77–83.)

by the fact that Dummett's critique is generally taken as having put it to rest. Be this as it may, this is no reason to give up trying to understand the nature of the modality argument.

On this score, two comments can be made at the outset. First, Dummett begins his defence of Frege thus:

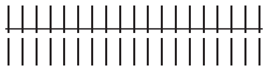
> The objection is readily answered. Frege invokes no modal notions: his definition is in terms of there *being* a suitable mapping. Waismann's objection can easily be reformulated as being that Frege owed us a criterion for the existence of relations, and that no such criterion can be framed without circularity.[12]

He then proceeds to show that such a criteria can be given without circularity, invoking in particular the axiom of choice for the (non-denumerably) infinite case, since one can *prove* with it the existence of one-one correlations between non-denumerably infinite sets. We have no qualms with this (at least for the moment), but one should note that Wittgenstein argues his point only for finite numbers: if one's wish is to understand the argument, it is better to restrict the discussion to this case, instead of attacking it in reference to a case it was not meant to cover. Secondly, this quotation shows that Dummett's objections are based on a reading of the modality argument as an 'ontological' argument about the *existence* of 'one-one correlations'; we think that this is incorrect and favour instead, following Boudewijn de Bruin, an 'epistemic' reading of it in terms of *knowledge* of 'one-one correlations'.[13]

The key to de Bruin's reformulation resides in noticing that Wittgenstein's own formulations are indeed in *epistemic* terms. It is not as if Wittgenstein was not wary of the ontological presuppositions of the Frege-Russell definition, i.e., about the *existence* of the 'one-one correlations' necessary for it to go through, as he frequently discusses them.[14] But his formulations of the modality argument are nearly always in terms of *knowledge* of 'one-one correlations', for example at the beginning of the following passage:

> Can I know there are as many apples as pears on this plate, without knowing how many? And what is meant by not knowing how many? And how can I find out how many? Surely by counting. It is obvious that you can discover that there are the same number by correlation, without counting the classes.

12. Dummett 1991, 148f.
13. See his de Bruin 1999; and de Bruin 2008.
14. For example, at AWL, 158, 161f., 164f.; PG 356; LFM, 162.



> In Russell's theory only an *actual* correlation can show the 'similarity' between the classes. Not the *possibility* of correlation, for this consists precisely in numerical equality. Indeed, the possibility must be an *internal* relation between the extensions of the concepts, but this internal relation is only given through the equality of the 2 numbers.[15]

With K standing for the usual operator from epistemic logic, de Bruin defines '*de re* knowledge' as knowledge that there is an object x such that one knows that it has the property P, or

$$\exists x\, K\mathrm{P}x$$

and '*de dicto* knowledge' as knowledge that there is an object x that has property P, or

$$K\exists x\, \mathrm{P}x$$

As is usually assumed, *de re* knowledge entails *de dicto* knowledge:

$$\exists x\, K\mathrm{P}x \rightarrow K\exists x\, \mathrm{P}x$$

So de Bruin introduces a notion of '*merely de dicto* knowledge', i.e., '*de dicto* but not *de re* knowledge':

$$K\exists x\, \mathrm{P}x\ \&\ \neg\exists x\, K\mathrm{P}x$$

These notions allow for the following reformulation of the modality argument. First, to draw a one-one correlation between the Fs and the Gs without any knowledge of 'how many' Fs and Gs there are gives us *merely de dicto* knowledge and can be reformulated as

$$K\exists n\, (\mathrm{N}x : \mathrm{F}x = n\ \&\ \mathrm{N}x : \mathrm{G}x = n)\ \&\ \neg\exists n\, K(\mathrm{N}x : \mathrm{F}x = n\ \&\ \mathrm{N}x : \mathrm{G}x = n)$$

This is the situation described above, where one can draw a one-one correlation for large numbers, and therefore know that there are equally many without knowing how many: one knows that there exists a number n which

15. PR, §118.

is the cardinality of F and G, but one does not know what number that is. Counting gives instead *de re* knowledge: one knows of some number n that it is the cardinality of the Fs and of the Gs, or

$$\exists n\, K(\mathrm{N}x : \mathrm{F}x = n \,\&\, \mathrm{N}x : \mathrm{G}x = n)$$

On the other hand, drawing a one-one correlation gives one *de re* knowledge of the one-one correlation, but this does not presuppose *de re* knowledge about sameness of cardinality. This notion of *de re* knowledge of the one-one correlation would also correspond to the notion of 'actual' one-one correlations in Wittgenstein's argument and it is the opposite to *merely de dicto* knowledge about a one-one correlation, which corresponds rather to the notion of 'possible' one-one correlations in Wittgenstein's argument.

With these notions at hand, one may indeed reformulate Wittgenstein's argument by simply pointing out that, according to him, *merely de dicto* knowledge of one-one correlation presupposes *de re* knowledge about sameness of cardinality; de Bruin has moreover argued that, under some constructivist principles about existence and knowledge, the modality argument is valid.[16] Again, we wish to steer clear of issues concerning the evaluation of the argument; we would like simply to ask for Wittgenstein's underlying arguments. Is there any reason why *merely de dicto* knowledge of one-one correlation would presuppose *de re* knowledge about sameness of cardinality? Why would one only *know* that a one-one correlation *can* be established only when one *already knows* that the two classes have the same number?

But asking this question is equivalent to asking: How could one establish a one-one correlation without counting? Let us thus suppose there are nine apples and nine oranges on a table. Of course, the purpose of a one-one

16. See de Bruin 2008, 365. These principles are: (1) for something to exist means that it be constructed; (2) every piece of knowledge must eventually rest on some constructive piece of knowledge; (3) there are precisely two independent ways to obtain knowledge about one-one correlation between two concepts, one involving one-one correlation, one involving cardinality. To assess the plausibility of attributing them to Wittgenstein would leave us far afield. At least this much shows that in order for his arguments to hold, Wittgenstein had to be committed to constructivist principles, a conclusion that the vast majority of his commentators have adamantly refused to draw; in despair they usually prefer to discount his philosophy of mathematics altogether. We should point out, however, that we assume here some equivalent to (3), which is that there are two independent ways to obtain knowledge about a one-one correlation, namely a direct way, by some form of subitization, and an indirect way, by counting.

correlation is *not* to find out 'how many' of these there are, it tells one only if there are 'as many' apples as there are oranges. Correlating them would mean something like putting an apple together with each orange, and when this procedure has come to an end, one can say that one now *knows* that there are as many apples as there are oranges. One can thus infer that 'The number of apples and oranges on this table is the same', without necessarily *knowing* what that number is. There is obviously no way one would know their number without resorting to counting, unless that number is small enough for one to take it in at a glance without any error.

It is a matter of human physiology, which is the topic of much research in psychology and neuroscience, that humans can recognize at a glance without failing numbers smaller than 4, and with occasional failure up to 7, but that, for higher numbers, they start counting. (One interesting point to make about subitizing is that the world's many abaci—Chinese, Japanese, Russian, etc.—are designed so that one can usually take in at a glance large numbers without subitizing numbers greater than five.) The process by which one immediately recognizes small numbers is called 'subitizing', from the Latin '*subitus*' or 'sudden'.[17] So, for very small numbers within the domain of subitization, one could recognize immediately the sameness of numbers. To circumvent such obvious limitations, one might arrange the sets of objects in a familiar pattern, e.g., two rows, so that one also immediately sees if they have the same numbers. Or one might in some cases, e.g., when these are figures on a sheet of paper, draw lines. One often finds the latter procedure in textbooks, to get the idea of a one-one correlation across to students.[18]

Wittgenstein was perfectly aware of these various criteria. For example, in section 115 of the *Big Typescript*, he wrote:

> Here incidentally there is a certain difficulty about the numerals (1), ((1) + 1), etc.: beyond a certain length we cannot distinguish them any further without counting the strokes, and so without translating the signs into different ones. "||||||||||" and "|||||||||||" cannot be distinguished in the same sense as 10 and 11, and so they aren't in the same sense distinct signs. The same thing

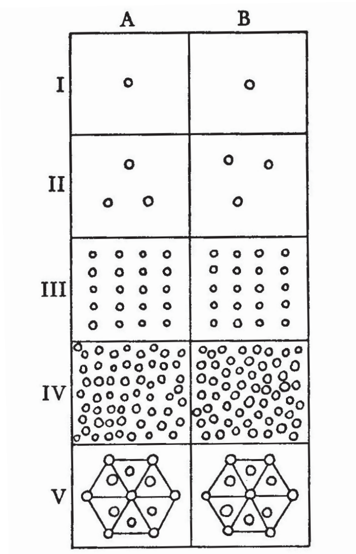
17. The term 'subitization' was introduced in Kaufman, Lord, Reese & Volkmann 1949. The idea that 'subitizing' and 'counting' are the result of two independent neural processes is still a matter of debate, but this is of no importance in the context of our discussion. For examples of contributions to this debate, see Simon & Vaishnavi 1996; Piazza *et al.* 2002; and Revkin *et al.* 2008.

18. See, e.g., Enderton 1977, 129 for one clear example.

could also happen incidentally in the decimal system (think of the numbers 1111111111 and 11111111111), and that is not without significance.[19]

And, in section 118 of the same typescript, he draws a variety of criteria, I to V:

Sameness of number, when it is a matter of a number of lines "that one can take it in a glance", is a different sameness from that which can only be established by counting the lines.



Different criteria for sameness of number. In I and II the number that one immediately recognizes; in III the criterion of correlation, in IV we have to count both groups; in V we recognize the same pattern.[20]

One will have recognize that subitization is involved in I and II.

These passages clearly show, therefore, that Wittgenstein was aware of the role played here by visual thinking.[21] He also saw that Russell & Whitehead implicitly rely on subitization:

19. BT, 398; PG, 330.

20. BT, 414; PG, 354.

21. One objection here would be to rule out our discussion by claiming that it amount to 'psychologism'. To show that it isn't would leave us to far afield, so we would like simply to refer, in the case of the modality argument, to de Bruin's explanations in de Bruin 2008, 366ff., and for the surveyability argument, below, to Marion 2011, 150f.

We actually say, "Well *this* is one and *this* is one." It is very important for the treatment of *Principia Mathematica* that there are classes whose numerical equality we can take in at a glance.[22]

And he did not condemn establishing one-one correlations by drawing lines either:

> [...] if asked whether *abc* and *def* could have different numbers, the answer is No, since these can be surveyed. Would you call it an experiment to correlate *abcd... w* and $\alpha\beta\gamma$ ... ω so as to *see* whether they have the same number? Would you say that you determine by experiment whether the number of numbers between 4 and 16 is the same as the number of those between 25 and 38? No, this is determined [...] using dashes or something similar.
>
> It is a pernicious prejudice to think that using dashes is an experiment and substraction a calculation. This is comparable to supposing a Euclidean proof by using drawing is inexact whereas by using words it is not.[23]

The study of visual thinking in mathematics or logic has been considered a forbidden zone since Frege. One should note that Wittgenstein clearly objects in this passage to the formalist tendency, perhaps exacerbated in the Hilbert school, to denigrate it. As it turns out, the study of visual thinking has recently become more respectable. Marcus Giaquinto, who has been one of the main contributors to this field, concluded a recent survey stating that:

> Visual thinking can occur as a non-superfluous part of thinking through a proof and it can at the same time be irreplaceable, in the sense that one could not think through the same proof by a process of thought in which the visual thinking is replaced by some thinking of a different kind.[24]

As we can see, Wittgenstein could not but agree with this: the modality argument, properly understood, is rather in line with Giaquinto's comment.[25] This, of course, goes against the grain of much of Wittgenstein scholarship, where commentators often premise their interpretation on a formalist stance, which is not open to discussion, so to rule out the sort of things we say here. But these passages are clear: Wittgenstein recognizes the role of visual thinking and faults Russell for misunderstanding it.

22. LFM, 164.
23. AWL, 158f.
24. Giaquinto 2008, 39f.; see also Giaquinto 2007.
25. One should note, however, that Giaquinto never discusses the point we are claiming Wittgenstein raised here, so we are not implying that there is a convergence between their ideas.

In a nutshell, his point is as follows. One-one correlations can be divided into two classes: a first class will contain those that are actual, in the sense that they are produced by one of the above criteria, subitization, pattern recognition, drawing lines, etc. It is a fact, however, that all these criteria will eventually *peter out when numbers grow large enough*. (To take an obvious example, one would not be able to correlate with any amount of certainty two sets of 3 million elements by drawing lines.) So there must be a second class which will comprise all the possible, non-actual one-one correlations. Wittgenstein's argument is thus, simply, that it is illegitimate to assume that what is sufficient for the first class, namely some form of visual recognition, is also sufficient for the second class. Thus the different criteria for producing an 'actual' one-one correlation (subitization, drawing lines, etc.) eventually peter out, and, once they have effectively come to an end, one is left with no other choice but to count, which would give *de re* knowledge about sameness of (cardinal) number. So merely *de dicto* knowledge of one-one correlation will presuppose, once other criteria become ineffective, *de re* knowledge. This is the answer to our question: Is there any reason why *merely de dicto* knowledge of one-one correlation would presuppose *de re* knowledge about sameness of cardinality?

3. *The Surveyability Argument*

The same point, we contend, lies at the heart of Wittgenstein's surveyability argument. We won't spend too much time reconstructing the argument and its implications, as one of us did it elsewhere.[26] The target here is what may be called Russell's version of mathematical 'explicativism', in particular a pair of theses explicitly framed by Mark Steiner:[27]

i) it is sufficient to understand proofs written in the system of *Principia Mathematica* in order to know all the truths of arithmetic that we know; and
ii) it is possible for us actually to come to know arithmetical truths by constructing logical proofs of them.[28]

26. See Marion 2011.
27. See Steiner 1975, 25.
28. Steiner talks here in terms of proofs, but our discussion below, with formulas (a)–(c), does not involve proofs.

In a well-known passage from *Introduction to Mathematical Philosophy*, Russell pointed out that the "primitive concepts" contained in Peano's axioms, '0', 'number', and 'successor', are "capable of an infinite number of different interpretations, all of which will satisfy the five primitive propositions" (Russell 1919, 7). Given one such interpretation, one obtains a 'progression', which he defined as a series with a beginning but endless and containing no repetition and no terms that cannot be reached from the beginning in a finite number of steps. There is indeed an infinity of such 'progressions' which will, like the series of natural numbers, satisfy Peano's axioms—it suffices for example to start any given series with a natural number other than 0. So Russell argued that in Peano's arithmetic[29] "there is nothing to enable us to distinguish between [...] different interpretations of his primitive ideas", while

> We want our numbers not merely to verify mathematical formulae, but to apply in the right way to common objects. We want to have ten fingers and two eyes and one nose. A system in which "1" meant 100, and "2" meant "101", and so on, might be all right for pure mathematics, but would not suit daily life. We want "0" and "number" and "successor" to have meanings which will give us the right allowance of fingers and eyes and nose. We have already some knowledge (though not sufficiently articulate or analytic) of what we mean by "1" and "2" and so on, and our use of numbers in arithmetic must conform to this knowledge.[30]

The idea here would be that an interpretation within the logical system of *Principia Mathematica* of Peano's axioms, provides a definite meaning to its basic number-theoretic concepts and that this interpretation would allow one to recover applications of arithmetic, i.e., that we have 'ten fingers and two eyes and one nose', etc. The very purpose of *Principia Mathematica* thus appears to be this:

iii) to set up an interpretation of Peano's axioms in order to provide a definite meaning to its primitive terms; and
iv) to recover ordinary applications of arithmetic.

The surveyability argument that Wittgenstein deploys against (i)–(iv) is easily stated by taking any ordinary number-theoretic equation, such as:[31]

29. The expression 'Peano's arithmetic' occurs at Russell 1919, 5, and in this original sense, it differs from today's frequent use of it as a name for first-order arithmetic.
30. Russell 1919, 9.
31. One of Wittgenstein's examples, ad RFM, III, §11.

(a) 27 + 16 = 43

According to Wittgenstein, this equation must have a counterpart in Russell & Whitehead's *Principia Mathematica*, of the form: [32]

(b) $(\exists!_{27}x(\mathrm{F}x)\ \exists!_{16}x(\mathrm{G}x)\ \&\ «x\ \neg(\mathrm{F}x\ \&\ \mathrm{G}x)) \rightarrow\ (\exists!_{43}x(\mathrm{F}x \vee \mathrm{G}x))$

Now, Russell's stance in (i)-(iv) amounts to an 'explicativist' claim of the sort '(a), and (a) *because of* (b)'.[33] Against this, Wittgenstein first noted that (b) must merely be an abbreviation of a longer formula with a total of 43 variables on each side of the sign for the conditional, or a formula with iterated '!' such as this:

(c) (∃!!!!!!!!!!!!!!!!!!!!!!!!!!!*x* (F*x*) ∃!!!!!!!!!!!!!!!!*x* (G*x*) & «*x* ¬(F*x* & G*x*)) →
(∃!!!*x* (F*x* ∨ G*x*))[34]

He could then easily point out that this unabbreviated formula is 'unsurveyable' in the sense that one cannot tell the precise number of iterations of '!' unless one starts counting them. Wittgenstein is simply relying here on the fact that human beings cannot tell at a glance (without counting) that there are 27 exclamation marks following the first existential quantifier of (c). This is but the same point (subitization) made above. And one should note that this is not an appeal to 'vagueness' (as most 'anti-realist' readers of Wittgenstein assume); there is nothing vague at all about the fact that there are 27 exclamation marks.[35]

It is thus hard to see what value there would be for '(a) *because of* (b)' given that, visually, the strings of '!' in (c) provide no certainty. Moreover, even for the abbreviated formula (b) one has to calculate in order to know what to write on the right-hand side of the conditional. Doing this would

32. One may wonder where Wittgenstein got formulas such as (b). The definition of addition in Part II, section B of *Principia Mathematica* is rather complicated because of the need to account for ambiguity of types and, as far as one can tell, there is no formula corresponding to (b). The closest is at *54.43:

$\vdash :. \alpha, \beta \in 1 . \supset : \alpha \cap \beta = \Lambda . \equiv . \alpha \cup \beta \in 2.$

See Marion 2011, 142f. for a discussion.

33. Again, for a justification of this claim, see Marion 2011, 143ff. & 152–155.

34. This notation is even suggested from AWL, 148 quoted below.

35. This is the point made with help of (13) in Marion 2011, 150. One should note that, as explained in that paper, this reading of the surveyability argument goes against decades of misunderstanding it in terms of 'strict finitism'.

presuppose the very knowledge of the number-theoretic equation (a) which is supposedly certified by (b). Therefore, rather than (a) being grounded on (b), it is (b) which requires *knowledge* of (a) (to see that it is true is an application of (a)). There appears, therefore, to be a circularity in Russell's attempt to ground number-theoretical equations on logic. This is not to say, however, that it is devoid of any interest, since it draws links between addition of natural numbers in a number-theoretic calculus on the one hand and the union of disjoint classes in a logical calculus on the other. It is just that this does not mean that the latter stands as *foundation* for the former, in accordance with the 'explicativist' claim '(a), and (a) *because of* (b)'.[36] In *Philosophical Remarks*, Wittgenstein put the matter this way:

> How can I know that |||||||||| and |||||||||| are the *same* sign? It isn't enough that they look *alike*. For having roughly the same *Gestalt* can't be what constitutes the identity of signs, but just their being the same in number.
>
> If you write (E |||||) etc. (E |||||||) etc. . ⊃ (E ||||||||||||) --- A you may be in doubt as to how I obtained the numerical sign in the right-hand bracket if I don't know that it is the result of adding the two left-hand signs. I believe that makes it clear that this expression is only an application of 5 + 7 = 12 but doesn't represent this equation itself.[37]

Of course, Wittgenstein is writing sloppily, but one recognizes in his formula A here a variant of (a), where the strokes stand for the strings of '!' in its unabbreviated version (c).

The link with the modality argument should be obvious and can be seen immediately by considering a possible objection. The point Wittgenstein is making with respect to the unabbreviated version of (b) is relying on the limits of subitization. One could try and obviate these limitations, without reverting to counting, by drawing lines between the occurrences of '!' on the left-hand side and those on the right-hand side of the conditional, thus putting them into a one-one correlation that shows that both sides have the same number. As the argument goes, the problem is that, without counting, one would still *not know* which number that is, and, further, that any attempt at producing a one-one correlation by means of drawing lines will peter out with larger numbers, for which one could never be certain if the procedure has been applied correctly or if a mistake has crept in.

On these points the surveyability argument bears more than a superficial resemblance with the modality argument. As a matter of fact it is

36 The idea is expressed, for example, at LFM, 260f.

37 PR, § 103.

so obvious that one wonders why it had remained hitherto unnoticed in the secondary literature. That Wittgenstein had the link between the two arguments always in mind can be seen from these following passages from, respectively, his 1933–34 and 1939 lectures:

> I shall now discuss the idea that "1+1=2" is an abbreviation of such statements as "If I have one apple in one hand, and another in the other, then I have two apples in both hands." In my notation this is: (E*1x*) *fx* (E*1x*) *gx* . (~∀*x*) (*fx* . *gx*)) ⊃ (E*2x*) *fx* ∨ *gx*.[38] Now is it true that "1 + 1 = 2" is an abbreviation of the underlined? […] To use a simple example:
>
> $$(\mathrm{E}^{||}x)fx.(\mathrm{E}^{|||}x)gx.\sim(\exists x)fx.gx.\supset.(\mathrm{E}^{|||||}x)fx \vee gx.$$
>
> Whether this is a tautology or not I decide by adding. Now does it correspond to 2 + 3 = 5? This implication says nothing (as it is either a tautology or a contradiction). […] What is queer about the functional notation (E*15x*) *fx* (E*27x*) *gx* . (~∀*x*) (*fx* . *gx*)) ⊃ (E*42x*) *fx* ∨ *gx* is that we never use it when we are asked to reckon how many apples we have. One has to do an addition before one knows what to write after the quantifier in the consequent.
>
> This leads directly to examination of Russell's and Frege's theory of the cardinal numbers, of which the fundamental notion is correlation.[39]

> Russell puts down (*xy*) (*uv*) ⊃ (*xyuv*) and proves this is a tautology. But suppose you had a greater number of terms—ten million on each side—what would you do? You say you will have to correlate them. Here—(*xy*) (*uv*) ⊃ (*xyuv*)— it looks as if there were just one way of correlating. But with the huge number—would you correlate them in the same way?
>
> Is there only one way of correlating them? If there are more, which is the logical way?—You can do any damn thing you please. If you really wanted to prove by Russell's calculus the addition of two big numbers, you would already had to know how to add, count, etc.[40]

Such passages are particularly enlightening, since Wittgenstein discusses the surveyability argument using the very premises of his modality argument, thus bringing to the fore the common presuppositions of both arguments.

---

38. We keep here Wittgenstein's odd notation, where '(E*1x*) *fx*' is short for '∃*x* F*x* & ¬∃*x*,*y* (F*x* & F*y*)'. Here, it is equivalent to '$\exists!_1 x$ F*x*'.

39. AWL, 147f. Incidentally, this passage is followed immediately by a statement of the modality argument.

40. LFM, 159. Again, the formula '(*xy*) (*uv*) ⊃ (*xyuv*)' is only a rough version of our formula (b) above.

4. *Concluding remarks*

In this paper we aimed for a better understanding of Wittgenstein's modality argument, on the basis of an epistemic reading of it, emphasizing the central role played by a basic idea about visualization in his critique of the Frege-Russell use of one-one correlation in order to define sameness of number. We then pointed out that the same idea about visualization is also the key to his surveyability argument. We believe that this was a necessary step towards a proper understanding of the latter, as well as a number of other topics, such as his non-extensional view of mathematics as based on numerical calculations and his understanding of proofs by mathematical induction. We also avoided throughout any assessment of the value of these arguments and in closing we would simply point out, with respect to the modality argument, that it would be wrong to judge Wittgenstein's intentions merely on the basis of it; his considered view is not nearly as negative as it looks like from reading the above. It suffices to see this that one considers section 118 of the *Big Typescript* (also reproduced as Part II, section 21 of *Philosophical Grammar*). The modality argument occurs in that section, but it is used merely to criticize Russell and its occurrence is actually followed by some developments aiming at (partly) recovering the biconditional between 'one-one correlation' and 'sameness of number', except that this is done in such a way that the result cannot serve for a definition of natural numbers of the Frege-Russell kind. We hope to explain how Wittgenstein proceeds in a further paper, but for the moment it suffices to say that Wittgenstein's position was not as 'radical' as one usually makes it to be: he did not reject Hume's principle as such, but merely tried to understand it in his own non-extensional idiom.

References

de Bruin, Boudewijn 1999: "Wittgenstein's Objections Against the Frege-Russell Definition of Number". In: *Proceedings of the International Wittgenstein Symposium*. Kirchberg/Wechsel: Austrian Wittgenstein Society, 109–113.

— 2008: "Wittgenstein on Circularity in the Frege-Russell Definition of Cardinal Number". *Philosophia Mathematica* III (6), 354–373.

Dummett, Michael 1991: *Frege. Philosophy of Mathematics*. London: Duckworth.

Frege, Gottlob 1980: *The Foundations of Arithmetic*. Oxford: Blackwell.

Giaquinto, Marcus 2007: *Visual Thinking in Mathematics. An Epistemological Study*. Oxford: Oxford University Press.
— 2008: “Visualizing in Mathematics”. In: Paolo Mancosu (ed.), *The Philosophy of Mathematical Practice*. Oxford: Clarendon Press, 22–42.
Goodstein, Reuben Louis 1951: *Constructive Formalism*. Leicester: Leicester University Press.
Isaacson, Daniel 1993: “On a Criticism by Wittgenstein of the Frege-Russell Definition of Natural Number”. *Journal of Symbolic Logic* 58, 764–765.
Kaufman, E., Lord, M., Reese, T. & Volkmann, J. 1949 : “The Discrimination of Visual Number”. *The American Journal of Psychology* 62, 498–525.
Landini, Gregory 2007: *Wittgenstein's Apprenticeship with Russell*. Cambridge: Cambridge University Press.
Marion, Mathieu 1991: *Quantification and Finitism. A Study in Wittgenstein's Philosophy of Mathematics*. Doctoral thesis, University of Oxford.
— 1998: *Wittgenstein, Finitism, and the Foundations of Mathematics*. Oxford: Clarendon Press.
— 2011: “Wittgenstein on the Surveyability of Proofs”. In: Marie McGinn & Oskari Kuusela (eds.), *The Oxford Handbook of Wittgenstein*. Oxford: Clarendon Press, 138–161.
Piazza, Manuela, Mechelli, Andrea, Butterworth, Brian & Price, Cathy 2002: “Are Subitizing and Counting Implemented as Separate or Functionally Overlapping Processes?”. *NeuroImage* 15, 435–446.
Simon, Tony & Vaishnavi, Sandeep 1996: “Subitizing and Counting Depend on Different Attentional Mechanisms: Evidence from Visual Enumeration in Afterimages”. *Perception and Psychophysics* 58, 915–926.
Steiner, Mark 1975: *Mathematical Knowledge*. Ithaca NY: Cornell University Press.
Revkin, Susannah, Piazza, Manuela, Izard, Véronique, Cohen, Laurent & Dehaene, Stanislas 2008: “Does Subitizing Reflect Numerical Estimation?”. *Psychological Science* 19, 607–614.
Russell, Bertrand 1919: *Introduction to Mathematical Philosophy*. London: Allen & Unwin.
Waismann, Friedrich 1951: *Introduction to Mathematical Thinking*. New York NY: Ungar.
— 1982: *Lectures on the Philosophy of Mathematics*. Amsterdam: Rodopi.

Grazer Philosophische Studien
89 (2014), 79–91.

WITTGENSTEIN ON FORMULAE

Esther RAMHARTER
University of Vienna

Summary
This paper discusses Wittgenstein's treatment of formulae. In particular, it will be shown that although Wittgenstein frequently investigates both *equations* (e.g. $x^2 + 1 = 0$, $\bigwedge_x x(x-1) = x^2 - x$) and *formulae in a narrow sense* (e.g. $\sin 2x$, $r^2\pi$), he rarely addresses the two together, let alone discusses the latter as parts of equations. Two issues will be raised for which this is of especial consequence. This sheds light on the change that Wittgenstein's understanding of the relation between formulae and generality underwent between the so-called middle phase and the *Remarks on the Foundations of Mathematics*.

1. *Formulae in Wittgenstein's writings—overview and terminology*[1]

The starting point of my deliberation is a simple observation: from as early as 1929 onwards equations occur frequently in Wittgenstein's notes, and he also deals with expressions like x^2 from the early 1930s. However he rarely discusses these together; indeed there are almost no instances in which he elaborates the role of formulae like x^2 as *parts of* equations.

To make my point clearly I must clarify my terminology. 'Formulae' will denote only mathematical formulae, not logical ones. By 'formula in the broader sense' I mean every (well-formed) mathematical expression containing a variable;[2] and among these formulae I distinguish between 'equations' and 'formulae in the narrower sense': An equation is a formula (in the broader sense) that 'substantially' contains an equality-sign and is used as an equation, for example: $x^2 + 1 = 0$ or $\bigwedge_x x(x-1) = x^2 - x$ used in some usual context. A 'formula in the narrower sense' is every formula

1. I am indebted to Pasquale Frascolla and Felix Mühlhölzer for averting the various errors that would otherwise be found in this paper.
2. Note that therefore an arithmetical equation like $25 \times 25 = 625$ is not an equation in my terminology.

(in the broader sense) that is not an equation, for example: sin $2x$, $r^2\pi$, … Furthermore y = sin $2x$ can also be considered as a formula in the narrower sense, if the 'y ='-part is irrelevant to what is meant.[3] A formula in the narrower sense is something that 'gives you an output, whenever you give it an input'.[4] It is important to note that this distinction between the two sorts of formulae is a distinction in use, not in form.[5]

In the paper I will first give an overview of Wittgenstein's use of formulae by presenting some more or less typical examples; then I will elaborate two issues in which Wittgenstein's separation of formulae in the narrower sense from equations, and in particular his disregard of formulae as parts of equations, are of especial consequence.

2. *Different sorts of formulae in Wittgenstein's writings*

2.1. *Equations*

Michael Wrigley (1993, 76ff.) argues that in the middle period Wittgenstein still holds—as an inheritance from the *Tractatus*—that mathematics consists of equations. It is not necessary to accept this strong claim in order to agree that, in thinking about mathematics, Wittgenstein at this time focuses on equations[6]—one of the first occurrences perhaps being the opening paragraph of *Wittgenstein and the Vienna Circle* (WVC). I quote only one example from the early 1930s[7], and give citations to passages that demonstrate how the topic of equations is manifested in his writings:

3. For example, if a mathematician says, 'Let us consider the function x^2' and writes down '$y = x^2$' (as is often done), this is in my terminology a formula in the narrower sense.

4. I do not use 'function' or 'term', because they already have a specified meaning in mathematics. Wittgenstein himself does not use the expression 'term' either.

5. This distinction is not a mathematical one. Furthermore, it does not matter if these definitions are vague in the sense that they admit cases in which it is not clear if a certain expression is to be called an equation or a formula in the narrower sense (or a formula at all), because all I aim to demonstrate are tendencies in Wittgenstein's thinking.

6. There is widespread literature on equations and equalities in Wittgenstein's writings, especially on propositions and equations: Frascolla 1994, 54–72; Marion 1998, ch. 6.1, on equations and generality: Marion 1998, ch. 4, on equality of numbers: Frascolla 1994, 15–21, 44–54; Marion 1998, ch. 2.1, 3.3, on equality and identity: Marion 1998, ch. 3.1.

7. I take the dates from Nedo 1993; although the dating of the *Bergen Electronic Edition* sometimes differs slightly, it does not touch upon any of my arguments.

Are all the variables in the following equations variables of the same kind?

$$x^2 + y^2 + 2xy = (x + y)^2$$
$$x^2 + 3x + 2 = 0$$
$$x^2 + ax + b = 0$$
$$x^2 + xy + z = 0$$

That depends on the use of the equations. – […] How do you prove the proposition 'No. 1 holds for all values of x and y' and how do you prove the proposition 'there are values of that satisfy No. 2?' There is no more and no less similarity between the senses of the two propositions than there is between the proofs. (PG, Part II,V §24)

Further occurrences of equations can be found in *Philosophical Remarks* (PR) XI §121, XII §130, XIII §150, XIV §164, §167f., §176) and *Philosophical Grammar* (PG) Part II, VI §§29–33, §§36ff. The context of the equations is in each instance the study of the role of proofs.[8]

Passages like the following can be said to be typical of the period between 1934 and *Remarks on the Foundations of Mathematics* (RFM); in these Wittgenstein struggles with the conception of meaning with respect to formulae, but in a way that combines mathematical and extra-mathematical content:

Betrachte die Ausdrucksform: 'Ich habe so viele Taschentücher, als $x^3 + 2x - 3 = 0$ ergibt', oder 'Die Zahl meiner Anzüge ist *n und* $n^2 + 2n + 2 = 0$.' Hat dieser Satz Sinn? (MS 116, 60 (1936)) [9]

See also, e.g., PG Part I, VI §84; MS 116, 119 (1937/38). I mention these only for the sake of completeness, but I will not comment on quotations of this sort any further. They can at least be seen as a hint that, in the intermediate phase between 1934 and 1937, formulae did not have a very definite place in Wittgenstein's thoughts.

In RFM Wittgenstein makes occasional, if somewhat sporadic, mention of equations: see RFM III §3, §47, §50, §52, V §39, VII §46.

8. Sometimes Wittgenstein uses formulae in discussions about real numbers, but there the formulae do not play any crucial role.

9. 'Consider the form of expression: "I have as many handkerchiefs as $x^3 + 2x - 3 = 0$ gives", or "The number of my suits is *n and* $n^2 + 2n + 2 = 0$." Does this proposition have sense?' (my translation, E. R.).

2.2 *Formulae in the narrower sense*

Very often we see Wittgenstein arguing against views like the following:

> Suppose the order to square a series of numbers is written in the form of a table, thus:
>
> | x | 1 | 2 | 3 |
> |---|---|---|---|
> | x^2 | | | |
> | | | | |
>
> It seems to us as if by understanding the order we add something to it, something that fills the gap between command and execution. (PG, part I, I §9; similar: PG, part I, IV §61, VII §86)

Whereas this mention of formulae in the context of considerations about rule-following uses an obvious example of a formula in the narrower sense, there are also occurrences of formulae in the narrower sense that do not appear so at first glance; in MS 118ff. from 1937, which was published as the first part of RFM, Wittgenstein writes:

> We use the expression: 'The steps are determined by the formula …' How is it used?—We may perhaps refer to the fact that people are brought by their education (training) so to use the formula $y = x^2$, that they all work out the same value for y when they substitute the same number for x. (RFM I §1 = PI §189)

$y = x^2$ is, when used as in this paragraph, not an equation in my terminology (see my distinctions at the beginning of this paper). More mentions of formulae in the narrower sense can be found in rule-following considerations, e.g. PI, part I, §151, §185, §226.

2.3 *Equations and formulae in the narrower sense occurring together*

Though, or precisely because, this category of occurrences of formulae is almost empty, it is the most important one for my purposes. Wittgenstein at one point talks, although rather unspecifically, about the relation between an equation and its parts:

> If we know the rules of elementary trigonometry, I can check the proposition $\sin 2x = 2 \sin x \cdot \cos x$, but not the proposition $\sin x = x - \frac{x^3}{3!} + \ldots$ but that means that the sine function of elementary trigonometry and that of higher trigonometry are *different* concepts. (PR XIII §151 = PG, part II, V §25)

But he lets his thought stop here and does not develop it any further.

One can also read the following quotation as saying something about equations and their sub-formulae:

> [W]e *call* formulae of a particular kind (with the appropriate method of use) 'formulae which determine a number y for a given value of x', and formulae of another kind, ones which 'do not determine the number y for a given value of x'. ($y = x^2 + 1$ would be one of the first kind, $y > x^2 + 1$, $y = x^2 \pm 1$, $y = x^2 + z$ of the second.) The proposition 'The formula … determines a number y' will then be a statement about the form of the formulae […]. (RFM I §1)

But no more detailed reflections upon the relation between equations and formulae in the narrower sense (constituting the former) can be found in Wittgenstein's writings.

3. *Consequences of the separation of equations from formulae in the narrower sense*

3.1 *Issue 1: The two meanings of an equation*

What Wittgenstein in 1929 said about proofs in mathematics and inductions (WVC, 33) is more easily accessible than most of what he would contribute to this topic later on. He addresses (in WVC, 33) many complementary components of this topic, only for his lines of thought to split asunder: On the one hand, he embarks on what are now called 'rule-following considerations' by means of formulae like x^2 (see above); on the other hand, he studies proofs in geometry (see e.g. WVC, 36; PR XII §131, XIII §152) and in arithmetic.

Geometrical and arithmetical propositions, Wittgenstein argues, cannot get their meaning by any other way than by their (method of) proof. (According to Pasquale Frascolla (1994, 125), in the time between 1929 and 1933 it was 'knowing how to prove the proposition' that gives the meaning; afterwards it was the proof itself.) Insofar as they are seen as building a syntactical system, not depending on 'exterior' entities of any sort, their meaning has to be given by their position in that system and this position is fixed by their derivation. In Wittgenstein's words:

> [I]f I can never verify the sense of a proposition completely, then I cannot have meant anything by that proposition either. Then the proposition signifies nothing whatsoever. In order to determine the sense of a proposition, I should

have to know a very specific procedure for when to count the proposition as verified. (WVC, 47)

> Its meaning must derive from its proof. What the proof proves is the meaning of the proposition (neither more nor less) … (PR XI §121, 144, Fn 1)

Compare also PG, part II, V §24 above.[10]

Things are different with respect to the way Wittgenstein treats formulae in the narrower sense. They *are* related to some entities external to the syntactical system to which the formulae belong; e.g. they are related to how we deal with numbers. Hence they gain their meaning from this relation—where this relation can either consist in following a rule—see PG, part I, I §9 above—or in induction (which is not mutually exclusive, as the quotation shows)[11]:

> [T]he letters are not at all the expression of generality, since generality in no way finds its expression in symbols; it shows itself in induction. A formula of algebra corresponds to an induction, but it does not express the induction for the reason that the latter is inexpressible.
>
> Thus if I wrote down:
>
> $$\begin{matrix} x \\ x^2 \end{matrix}$$
>
> I should not yet know how to apply this rule; I have not as it were expressed the general rule, I have once again only formed a certain configuration of letters; for x is just as much an individual sign as 1, 2, 3. […] Generality shows itself in application. I have to read this generality into the configuration. But it is neither easier nor more difficult to recognize the general rule in the expression
>
> $$\begin{matrix} x \\ x^2 \end{matrix}$$
>
> than it was previously in the case where I recognized it from the individual numbers. (WVC, 154)

In a sense, the two lines of thought meet again in two extensive discussions of inductions in PR IX and PG, part II, VI. Let us consider now what happens when these two lines of thought are applied to algebraic propositions, which are equations:

10. For the relation between mathematical proposition and proof see: Frascolla 1994, 54–72; Mühlhölzer 2010, especially ch. II.2 and II.6; Potter 2011, 127ff.

11. See also PR XII §142.

> The verification is not *one* token [Anzeichen] of the truth, it is *the* sense of the proposition. [...]
> [T]hrough the induction [the algebraic propositions] gain their sense, not their truth. (PR XIV §166f.)

We are confronted with two meanings for algebraic formulae. And the following statement, written between the preceding two, makes the situation even worse:

> An induction doesn't prove the algebraic proposition, since only an equation can prove an equation.

In other words: only proof gives the sense, induction is not a proof, induction gives the sense[12]—Wittgenstein literally contradicts himself within a few lines.[13]

Wittgenstein sometimes calls induction a proof (see WVC, 135; RFM III §54), and sometimes refuses to do so (see WVC, 33; PR XIV §167). I will not discuss his arguments here, but merely remark that he tends to accept it as a proof where he is interested in proof, and to deny that it is a proof where he is interested in meaning.[14] Whether induction is a proof or not, we still face the fact that algebraic equations possess two meanings; and this is obviously the result of separating the discussions of formulae in the narrower sense from those of equations.[15] But one must be more precise about what is happening here: Induction or the relation to numbers —being initially studied in connection with formulae in the narrower sense —and proof—initially studied in connection with propositions (equations among them)—merge into 'the meaning' of algebraic equations. Thereby the equations obviously survive, but the formulae in the

12. Potter 2011 also states that only proof is what gives the sense and that induction gives the sense, but then concludes (127) that the sense of an algebraic equation is its inductive proof.

13. Note that these sentences have always been grouped together, the contradiction is therefore not the product of some later editorial work.

14. The core of his argument against induction being a proof is: 'If one regards the proof [of a formula $A(c)$ by induction, E.R.] as being of the same sort as the derivation of $(x + y)^2 = x^2 + 2xy + y^2$, then it proves the proposition "$A(c + 1)$" on the assumption "$A(c)$", and so of the proposition I really want to prove.' (PR XIV §164) Induction does not look or work like a derivation. Literature on induction (quantification over infinite domains): Frascolla 1994, 72–85; Lampert 2008; Marion 1998, ch. 4, Mühlhölzer 2008, 124f; Mühlhölzer 2010, 405–416; Shanker 1987.

15. To anticipate a possible objection: In the quoted passage WVC, 154 Wittgenstein *does* talk about equalities also, but about arithmetical, not about algebraic ones, and hence not about equations in my terminology.

narrower sense disappear: Wittgenstein does not pay any attention to the parts of the algebraic equations, which are formulae in the narrower sense. I will not follow the consequences of this here, but it will be important in Issue 2.

Wittgenstein would of course not find it acceptable to simply say that there are two meanings, but then neither does he explicitly comment on the relation between these two meanings. He *does*, however, explain the relation between induction and proof: They do different things, but are correlated. (See, e.g., PG VI §38: Accepting something like the associative law 'as a rule for a calculation with letters […] brings this calculus in a certain sense into unison with the calculus of the cardinal numbers.') The problem therefore has to be genuinely one of meaning: In the early 1930s Wittgenstein still wanted to identify *one* 'thing' (respectively) as the meaning of whatever has a meaning at all—an idea probably still attributable to the influence of Frege.[16] In the context of equations—and hence propositions—the one, uniquely determined 'thing' is the proof (note his insistence that different proofs make up different propositions, see WVC, 109); in the context of formulae in the narrower sense it is just the way we proceed, the way we handle the numbers, and/or induction. Where these two contexts overlap, the stated problem arises.

In RFM, as we know, this problem was solved by broadening the concept of meaning, allowing different components to contribute to the meaning. Wittgenstein states 'What is the criterion for the way a formula is meant? Presumably the way we always use it […]' (RFM I §2), and RFM I §1 lists several examples of such a use.[17]

In sum, one could say that the necessity to broaden the concept of the meaning of formulae was already there in the early 1930s, but it took the form of a contradiction, because formulae in the narrower sense and mathematical propositions each have *their* meaning, but in the case of algebraic equations these two meanings 'clash' and should be *one*.

16. Compare WVC, 135: "[T]he proposition is related to the proof as a sign is to the thing signified. The proposition is a name for the induction."

17. Wrigley (1993, 82) states: 'Clearly, then [in RFM], Wittgenstein no longer regards the proof/meaning thesis as the *whole* truth about what determines meaning in mathematics'. Even in the context of equations the view of meaning being exclusively proof (or method of proof) cannot be maintained so easily anymore. Additionally, Wittgenstein is generally, and for different reasons, less focused on propositions or equations in the late period (see Frascolla 1994, 127).

3.2 *Issue 2: Is there such a thing as a mathematical problem?*

The question of how mathematical problems are possible can be found in the middle as well as in the late period, but it takes different forms.

In the middle period understanding a proposition is identified with knowing how to verify it. This conception, at first glance, leaves space for mathematical problems. As Frascolla showed, however, this position led to its self-destruction (Frascolla 1994, 114), and, furthermore, even if it can be established as a coherent position it does not allow an explanation of what we *actually* mean by a 'mathematical problem': a proposition of which we do *not* know how to prove.

In the late period Wittgenstein sometimes claims that meaning is proof; if it is the case that we cannot understand the proposition if we do not know the proof, then this leaves even less space for the possibility of mathematical problems than the view of the middle period.

But, in spite of these conceptions that closely link the proposition to its proof, Wittgenstein states that his explanations 'mustn't wipe out the existence of mathematical problems' (PR XIII §148). As Felix Mühlhölzer 2010, 354 puts it: Wittgenstein wants to maintain the tension between the proof being the meaning of a mathematical proposition and a certain independence of the proposition from the proof. (Cf. e.g. RFM V §42.)

Wittgenstein's statements concerning what can be positively said about mathematical problems are quite non-committal and indefinite:[18]

> The difficult mathematical problems are those for whose solution we don't yet possess a *written* system. The mathematician who is looking for a solution then has a system in some sort of psychic symbolism, in images, 'in his head', and endeavours to get it down on the paper. (PR XIII §151)

> [T]he mathematicians are not *completely* blank and helpless when they are confronted by this proposition [Fermat's Last Theorem]. After all, they try certain methods of proving it; and so far as they try methods, so far do they understand the proposition. – But is that correct? Don't they understand it just as completely as one can possibly understand it? (RFM VI §13)

What if the proposition turns out to be wrong, Wittgenstein continues. He then argues back and forth and arrives at: "Understanding' is a vague concept.'

18. In PG, part II, III §11 Wittgenstein mentions the mathematician's 'instinct'. See also PG, part II, V §22.

Wittgenstein acknowledges that the mathematician possesses some sort of approach to unproved mathematical propositions, particularly ones in the form of an equation. But, for a considerable time, an equation constitutes an opaque entity for him (an exception is PR XIII § 151—see above). He almost ignores the fact that we know a lot about its parts, devoting particularly little thought to *what* we may know about these parts. Parts of equations, formulae in the narrower sense, are already (syntactically and semantically) interrelated with other elements of the same or a different system: There are restrictions on how to build them (e.g. you must not build $1/\sin x$ out of 1 and $\sin x$ or a/b out of a and b without further restrictions), rules about how to transform them, they belong to certain classes and not to others (e. g., formulae which determine a number y for a given value of x, and formulae of another kind—see RFM I §1), … And each of the subformulae of an equation is already related to numbers (not only the equation as a whole).

It is just as with parts of a jigsaw puzzle, of which we do not know if they build a certain picture. Now, one can argue that we indeed do not 'understand' a jigsaw puzzle if we do not know the result – which is what Wittgenstein says about mathematical propositions. But then one has implicitly admitted that with mathematical propositions it is the same as with certain non-mathematical propositions: We have some knowledge about them, but we do not know everything; we may know enough or not enough for a certain purpose. And this is indeed what Wittgenstein finally arrives at:

> Thus it is as if the proof did not determine the sense of the proposition proved; and yet as if it did determine it. But isn't it like that with any verification of any proposition? (RFM VI §10)

And he goes on:

> A proof of the proposition locates it in the whole system of calculations. And its position therein can now be described in more than one way. (RFM VI §11)

(Wittgenstein now concedes that an equation can have different proofs.) I think it would be quite in line with this view to say that, *in a sense*, the position of an equation can be given even without the proof. In the case of Wittgenstein's example $\sin 2x = 2 \sin x \cdot \cos x$ one could say: The formula $\sin 2x$ has its position within a given calculus, and so have $\sin x$ and $\cos x$ and hence $2 \sin x \cdot \cos x$. The position of the equation within the system is fixed by saying that it is what we get when we put an equality-sign between the former and the latter. But *in another sense* the position in the system is not fixed as long as we do not know the proof—which makes mathematical problems possible. I of course do not claim that this is in anyway a new thesis—it is merely how mathematicians would describe the matter—, but the question is whether it can be developed out of Wittgenstein's investigations. So far I have argued that it is compatible with Wittgenstein's remarks; now I wish to add an argument for why, in a way, it is a consequence of them.

The suspicion that mathematical problems could not possibly exist is based on a particular understanding of mathematical propositions: an equation (e.g.) gains its meaning as a whole from its proof—which consists of equations—, not via its parts. However the transformation rules include the use of formulae in the narrower sense (such as multiplying the right and the left side of the equation by n^2), and because therefore the proof involves formulae in the narrower sense we cannot have an understanding of a proposition and its proof without already having an understanding of formulae in the narrower sense. Here one could object that we need not have an *understanding* of the formulae involved in the transformations – we just use them to transform one equation into another. However this cannot be maintained from Wittgenstein's point of view, as one of his examples shows: he repeatedly (RFM III §78, §85f, VII §14) wonders if or under what circumstances it could make sense to divide or multiply an equation by $(n-n)$. Here it is clear that $(n-n)$ must have some sort of meaning, before it becomes a part of the equation or is 'applied' to the equation, otherwise the discussion would be pointless. Therefore that the 'position [of the proposition can] be described in more than one way' has to imply that this position is partially—or in a certain respect—determined by formulae in the narrower sense.

I conclude: Swayed by his prior interest in mathematical propositions, Wittgenstein does not pay a great deal of attention to the subformulae of equations[19]—in contrast to what he had always done with respect to empirical propositions (previously in the *Tractatus*)—and therefore (re-)constructing an adequate place for the 'mathematical problem' leads the reader on a winding path through remarks scattered throughout Wittgenstein's notes.[20]

19. I find it amusing that Wittgenstein studies the quite extraordinary case of equations as parts of ordinary propositions (see MS 116, 60 above), but not the parts of equations.

20. To study from this perspective the relation between the mathematical problem and the surprising within mathematics (see Floyd 2012), for example, would be an interesting task.

Finally, I offer an outlook: The context principle deserves attention as a backdrop to my deliberations. Peter Sullivan, in discussing Frege and the *Tractatus*, formulated an abstract version of the context principle of sense: 'the sense of a subsentential expression consists in, and is exhausted by, its systematic contribution to the thoughts expressed by propositions in which it figures' (Sullivan 2001, 75). As regards mathematics in Wittgenstein's middle and later period, several modifications and ideas suggest themselves, e.g.: With respect to equations or mathematical propositions in general and their proofs, a shift becomes necessary: in some of Wittgenstein's remarks the proposition as a whole plays the 'sub'-role, the proof being its sense-giving context (furthermore, the formulae in the narrower sense, which we usually tend to see as the components of the equations, become the context of the equations—see above). Further, ignoring the contribution of the subformulae of equations to the meaning would create remarkable entities: a mathematical proposition would be a proposition that is not a context of anything.[21]Taking my observations in this paper as a starting point, it seems to me that a (re-)examination of what serves as a 'context' in Wittgenstein's work on mathematics, and of the effects on his understanding of 'meaning' in general (not restricted to mathematics), is a promising line of enquiry.

References

Floyd, Juliet 2012: "*Das Überraschende*: Wittgenstein on the Surprising in Mathematics". In: Jon Ellis & Daniel Guevara (eds.), *Wittgenstein and the Philosophy of Mind*. Oxford – New York: Oxford University Press, 225–258.

Frascolla, Pasquale 1994: *Wittgenstein's Philosophy of Mathematics*. London: Routledge.

Lampert, Timm 2008: "Wittgenstein on the Infinity of Primes". *History and Philosophy of Logic* 29 (3), 272–303.

Marion, Mathieu 1998: *Wittgenstein, Finitism, and the Foundations of Mathematics*. Oxford: Clarendon Press.

Mühlhölzer, Felix 2008: "Wittgenstein und der Formalismus". In: Matthias Kroß (ed.), *"Ein Netz von Normen", Wittgenstein und die Mathematik*. Berlin: Parerga, 107–148.

21. I am well aware that much more would have to be said about whether a mathematical proposition is a proposition.

Mühlhölzer, Felix 2010: *Braucht die Mathematik eine Grundlegung? Ein Kommentar des Teils III von Wittgensteins* Bemerkungen über die Grundlagen der Mathematik. Frankfurt am Main: Klostermann.

Nedo, Michael (ed.) 1993: *Ludwig Wittgenstein. Wiener Ausgabe. Einführung/Introduction*. Wien-New York: Springer.

Potter, Michael 2011: "Wittgenstein on Mathematics". In: Oskari Kuusela & Marie McGinn (eds.), *The Oxford Handbook on Wittgenstein*. Oxford: Oxford University Press, 122–137.

Shanker, Stuart 1987: "Wittgenstein's Remarks on the Significance of Gödel's Theorem". In: Stuart Shanker (ed.), *Gödel's Theorem in Focus*. London: Croom Helm, 155–256.

Sullivan, Peter 2001: "Wittgenstein's Context Principle". In: Wilhelm Vossenkuhl (ed.), *Ludwig Wittgenstein. Tractatus Logico-Philosophicus*. Berlin: Akademie Verlag, 65–110.

Wrigley, Michael 1993: "The Continuity of Wittgenstein's Philosophy of Mathematics". In: Klaus Puhl (ed.), *Wittgenstein's Philosophy of Mathematics, Proceedings of the 15th International Wittgenstein Symposium, Part 2*. Wien: Hölder-Pichler-Tempsky, 73–84.

Grazer Philosophische Studien
89 (2014), 93–108.

BROUWER *VERSUS* WITTGENSTEIN ON THE INFINITE AND THE LAW OF EXCLUDED MIDDLE

Ian RUMFITT
Birkbeck College, University of London

Summary

Wittgenstein and Brouwer were agreed that some of the higher mathematics of their day rested upon a projection into the infinite of methods that legitimately apply only within finite domains. In this paper I compare and assess the different treatments the two philosophers give of problematic cases involving infinity. For Brouwer, certain claims about infinite sequences provide exceptions to the law of excluded middle; while Wittgenstein argues that the same claims are without sense, since for him the law of excluded middle is a criterion of being a proposition. I end the paper by outlining how the intuitionist might respond to Wittgenstein's arguments.

According to Herbert Feigl, who was with him on the day, Wittgenstein was provoked into returning to philosophy by hearing L. E. J. Brouwer's lecture, 'Mathematik, Wissenschaft und Sprache', in Vienna on 10 March 1928 (see the quotation from Feigl in Pitcher 1964, 8*n*). While Wittgenstein's later writings reject several central Brouwerian theses, a comparison between these thinkers is instructive. As I hope this paper will show, Wittgenstein accepts one of Brouwer's key negative contentions—namely, that some of the higher mathematics of their day rests upon an illegitimate projection into the infinite of methods that properly apply only within finite domains. While they differ over the remedy, agreement on that negative point and Wittgenstein's close engagement with Brouwer's positive theory belie the widespread view—inspired by a notorious *obiter dictum* in the transcript of a 1939 lecture—that, for Wittgenstein, 'Intuitionism is all bosh—entirely' (LFM 237).[1]

1. The account of intuitionism that directly precedes this dictum in the lecture notes (which were taken down by some students) is in any case eccentric.

1. *The intuitionists on infinity*

Nowadays, under the influence of the late Sir Michael Dummett, we are apt to associate the intuitionist critique of classical mathematics and logic with the adoption of verificationist semantic theories, in which the meaning of a declarative sentence (henceforth, a statement) is given by specifying the conditions in which a speaker would be entitled to assert it, rather than by specifying the conditions under which it would be true. It is important to set these associations aside in reading the early intuitionists, for the founding fathers of the school were not verificationists. In a paper of 1923, Brouwer wrote that 'a complete empirical corroboration of the inferences drawn [about the "world of perception"] is usually materially excluded a priori and there cannot be any question of even a partial corroboration in the case of (juridical and other) inferences about the past' (Brouwer 1923, 336). A verificationist would conclude from that claim that talk about the past is meaningless; Brouwer, though, expressly holds that it is meaningful. Indeed, he allows that the laws of classical logic, including Excluded Middle, may validly be applied in reasoning about the world of perception, as long as we are able to think of the 'objects and mechanisms of [that] world … as (*possibly partly unknown*) finite discrete systems' (*ibid.*, emphasis in the original). More exactly, it is the possibility of projecting 'a finite discrete system upon the objects in question' that is the 'condition of the applicability' of Excluded Middle to judgements concerning those objects. We see here a fundamental difference between Brouwer and Dummett. For Dummett, the basic mistake of the classical mathematicians is that they apply a realist or truth-conditional semantic theory to the language of mathematics. For Brouwer, by contrast, their error was to apply distinctively classical logical rules 'even in the mathematics of infinite systems', where the rules' condition of applicability does not obtain. A. N. Kolmogorov, another pioneer of intuitionism, agreed with Brouwer. He understood Brouwer's writing to have 'revealed that it is illegitimate to use the principle of excluded middle in the domain of transfinite argument' (Kolmogorov 1925, 416).

As Brouwer's reference to 'infinite systems' implies, the early intuitionists did not impugn as unintelligible expressions, such as 'the sequence of natural numbers', that purport to designate infinite mathematical structures. They did, however, claim that talk about such structures, if it makes sense at all, is disguised talk about the mathematical principles that characterize them. Thus, to say that the natural number sequence has a property is to say that the property in question is entailed by the laws of Heyting Arithmetic, these laws (the intuitionistic analogue of the Peano Postulates) being the principles that characterize that structure. This marks a fundamental contrast with the finite case. A finite structure might be characterized by certain mathematical principles but, even when it is so characterized, it may still have properties that are not entailed by the principles. As one might put it, in the finite case the *extension* of certain mathematical principles will have mathematical properties over and above those consequent on the principles themselves. According to the intuitionist, this is conceptually impossible in the infinite case. A finite initial segment of an infinite sequence may have properties over and above those entailed by the principles that generate the sequence. But if we speak of the infinite sequence as a whole, we must be referring (perhaps elliptically) to the generating principles themselves. For the intuitionist, one might say, infinite structures cannot be conceived purely extensionally. So to conceive them is illegitimately to project into the infinite a notion that only makes sense in the finite case.

Wittgenstein understood and heeded Brouwer's warning not to treat infinite collections as though they were large finite ones. In §19 of Part V of the *Remarks on the Foundations of Mathematics* (RFM), which its editors date to between 1942 and 1944, he asks:

> Isn't it like this? The concepts of infinite decimals in mathematical propositions are not concepts of series, but of the unlimited technique of expansion of series.
>
> We learn an endless technique: that is to say, something is done for us first, and then we do it; we are told rules and we do exercises in following them; perhaps some expression like 'and so on *ad inf.*' is also used, but what is in question is not some gigantic extension (278f.).

A little later, in §36, he says:

> Our difficulty really already begins with the infinite straight line; although we learn even as children that a straight line has no end, and I do not know that this idea has ever given anyone any difficulty … But the straight line is a *law* for producing further (290).

Remarks such as these—which are typical of Part V of RFM—nicely express Brouwer's basic objection to the conception of the infinite that prevailed in his day and still prevails in ours. While Wittgenstein and Brouwer differ over the best prophylactic against this popular misconception, they are at one in perceiving a deep problem in the standard view of the infinite, and as such they are allies against the majority of mathematicians.

2. *Brouwer against the Law of Excluded Middle*

According to classical logic, we are entitled to assert $\ulcorner A \vee \neg A \urcorner$ no matter what meaningful statement A might be. Brouwer argues, though, that there are meaningful mathematical statements A for which an assertion of $\ulcorner A \vee \neg A \urcorner$ conflicts with a correct view of the infinite. Accordingly, a correct view of the infinite forces us to revise classical logic. In particular, it forces us to restrict the Law of Excluded Middle.[2] Since this revisionist claim is one that Wittgenstein rejects, it will be worth setting out Brouwer's grounds for it carefully.

In the Vienna lecture that Wittgenstein heard, Brouwer introduced the notion of a *Pendelzahl*—a pendulum number or (as he Englished his term) a 'binary oscillatory shrinking number'. He then argued that we are not entitled to assert that such a number is either identical with or distinct from zero (Brouwer 1928, 1183).[3] Wittgenstein evidently remembered the example, for in the *Philosophical Remarks* of 1929–31 he wrote:

> Brouwer is right when he says that the properties of his *Pendelzahl* are incompatible with the law of the excluded middle. But, saying this doesn't reveal a peculiarity of propositions about infinite aggregates. Rather, it is based on the fact that logic presupposes that it cannot be *a priori*—i.e. logically—impossible to tell whether a proposition is true or false. For, if the question of the truth or falsity of a proposition is *a priori* undecidable, the consequence is that the proposition loses its sense, and the consequence of this is precisely that the propositions of logic lose their validity for it (PR 210).

In the light of the developments initiated by Gödel's great paper of 1931, philosophers and logicians will demand a great deal of argument before they can be persuaded to take seriously, let alone accept, Wittgenstein's

---

2. The restriction consists in our not being entitled to assert certain instances of Excluded Middle. For the intuitionist, no such instance is false, i.e. has a true negation. For in intuitionistic logic $\neg(A \vee \neg A)$ entails the patently contradictory $\neg A \wedge \neg\neg A$.

3. Brouwer actually wrote that 'this binary oscillatory shrinking number is neither equal to zero, nor different from it—in violation of the principle of the excluded middle'. As Ewald remarks (1996, 1183, *n.t*), these words need to be read charitably if Brouwer is not to find himself embroiled in the contradiction identified in the previous footnote.

claim that undecidable propositions lack sense.[4] For present purposes, though, we need not address that large issue. For in other writings from the 1920s, Brouwer presents rather simpler instances of Excluded Middle which (as he thinks) we are not entitled to assert and to which Wittgenstein responded with a detailed analysis, not a sweeping denial of sense to all undecidable statements.

Brouwer presents the sort of case I have in mind in subtly different ways in different places, but the exposition in his 1923 lecture and paper, 'On the significance of the Principle of Excluded Middle in mathematics', is characteristic. He begins §2 of that paper by identifying two 'fundamental properties'—propositions which are foundational for the current 'mathematics of infinity' and which follow from Excluded Middle. The second of these propositions is that every mathematical species is either finite or infinite. He then presents an example to show that this latter proposition are incorrect:

> Let d_v be the vth digit to the right of the decimal point in the decimal expansion of π, and let $m = k_n$ if, as the decimal expansion of π is progressively written, it happens at d_m for the nth time that the segment $d_m d_{m+1} \dots d_{m+9}$ of this decimal expansion forms the sequence 0123456789 … That the second fundamental property is incorrect is seen from the example provided by the species of the positive integers k_n defined above (Brouwer 1923, 337).

In other words, we cannot assert that the species of integers k_n is either finite or infinite.

Brouwer's species is surely well defined. This is because, for any integers m and n, there is a finite procedure that decides whether $m = k_n$. For suppose we wish to find out whether 538,763 = k_2. To do this, it suffices to calculate π to the first 538,772 decimal places. If the last 10 digits in the expansion are 0123456789, and if that segment occurs precisely once earlier in the expansion, then 538,763 = k_2; otherwise, it is not. A Turing machine could be programmed to apply this test, and it would report an answer in a finite time. For these reasons, it seems clear that Brouwer has identified a mathematically well-defined species of integers.

Why, though, does Brouwer maintain that we cannot assert that the species is either finite or infinite? While he is not fully explicit, I think

4. For Gödel—as, I take it, for Wittgenstein in PR—a statement is decidable (with respect to a theory T) if and only if either it or its negation is deducible from T (Gödel 1931, 597). A statement may be decidable in this sense with respect to the whole currently corpus of accepted mathematical theory even though there is no decision procedure for determining its truth-value.

the reason is clear. The species of k_n's is finite if and only if there are only finitely many segments of the form 0123456789 in the decimal expansion of π; and it is infinite if and only if there are infinitely many such segments. Accordingly, if we were entitled to assert 'Brouwer's species is either finite or infinite', we would also be entitled to assert 'Either (1) there are only finitely many segments 0123456789 in the decimal expansion of π or (2) there are infinitely many such segments'. Given Brouwer's strictures on the meaning of talk about the infinite, however, we are not entitled to assert that either (1) or (2) obtains. According to those strictures, a statement about an infinite sequence must be cashed out in terms of the principle or rule that generates the sequence. Given that, alternative (1) can only mean that the rule for expanding π entails that there are only finitely many segments of the form 0123456789 in the expansion. *Pari passu*, alternative (2) can only mean that the rule entails that no bound can be set on the number of such segments. In our present state of knowledge, we are not entitled to assert that either (1) or (2) obtains. Of course, our knowledge might expand in such a way that we become entitled to assert this. For example, a mathematician might prove, on the basis of the rule for expanding π, that there could be at most three occurrences of the segment 0123456789 in its decimal expansion; we would then know that alternative (1) obtains. In our present state of knowledge, however, we are not entitled to assert that either (1) or (2) obtains, and so we cannot assert that Brouwer's species is either finite or infinite.

In fact, it will help to work with a slightly simpler example. At the time of writing, π has been calculated to the first ten trillion (10^{13}) digits. I do not know whether those ten trillion digits include a segment 0123456789, but let us suppose that they do not. (If they do, one could easily change the designated segment to one that does not appear in the largest expansion of π that we currently have.) Let us now consider the statement 'Either Brouwer's species of k_n's is inhabited or it is not'. Given our supposition, we are not entitled to assert this instance of Excluded Middle. Brouwer's species is inhabited if and only if the segment 0123456789 occurs somewhere in the decimal expansion of π, and it is uninhabited (i.e. empty) if and only if no such segment occurs. So we would be entitled to assert 'Either Brouwer's species is inhabited or it is not' only if we were also entitled to assert 'Either 0123456789 occurs somewhere in the expansion of π or it does not'. Given Brouwer's strictures on what statements about the infinite can mean, the latter instance of Excluded Middle means 'Either (1) the rule for expanding π entails that the segment 0123456789 occurs somewhere in the expansion, or (2) the rule for expanding π entails that no such segment occurs anywhere'. In our current state of knowledge, we are not entitled to assert this disjunction. As before, this might change. In calculating π to the first twenty trillion digits, we might find a segment 0123456789; we would then know that alternative (1) obtains. Equally, a mathematician might prove that (2) obtains. In our present state of knowledge, though, we cannot assert that either (1) or (2) obtains; hence we cannot assert that Brouwer's species is either inhabited or not.

3. *Wittgenstein on unassertible instances of Excluded Middle*

I have switched to this simpler example in order to bring Wittgenstein back into the story, for a central question in Part V of RFM is precisely whether we are always entitled to assert that a given segment of digits either is or is not to be found somewhere in the decimal expansion of π. The fact that Wittgenstein focuses so intently on this question suggests forcibly that he had studied either the 1923 lecture from which I have quoted, or one of the other papers from the early 1920s in which Brouwer uses the same technique to cast doubt on the Law of Excluded Middle. At any rate, his focus surely refutes the hypothesis that, on Wittgenstein's considered view, intuitionism is 'bosh'—if that means that it is so confused as not to be worth discussing. As we have seen, the question Wittgenstein addresses is central to the intuitionist's critique of classical mathematics, and the paragraphs—from §9 to §23 of Part V—in which he develops his answer to it constitute one of the most sustained passages of argument in the whole of the *Remarks*. In gauging Wittgenstein's attitude to intuitionism, these facts must carry greater weight than a stray remark in a lecture.

In the *Philosophical Remarks* of 1929-31, and in his lectures of 1932-5 (AWL), Wittgenstein agrees with Brouwer that we are not entitled to assert certain instances of Excluded Middle. But they offer different diagnoses of why we are not always entitled to make such assertions. On Wittgenstein's view, the unassertible cases are not properly regarded as exceptions to the Law. Rather, statements like 'The segment 0123456789 occurs somewhere in the decimal expansion of π' do not qualify as meaningful propositions. Since the laws of logic apply only to propositions, these statements simply fall outside their ambit:

> I need hardly say that where the law of excluded middle doesn't apply, no other law of logic applies either, because in that case we aren't dealing with propositions of mathematics. (Against Weyl and Brouwer.) (PR 176)

The intuitionists, then, were misguided in seeking a non-classical logic to regulate inferences involving undecidable statements about the infinite: since such statements fail to qualify as propositions, they have no logic. Similarly, in his lectures of the early 1930s, Wittgenstein maintained that a willingness to take $\ulcorner A \vee \neg A \urcorner$ to be a tautology partly defines what it is for A to be a proposition. 'This pattern occurs somewhere in this expansion' is an example of a grammatically well-formed statement that *seems* to qualify as a proposition but in fact does not (AWL 140). On Brouwer's account, we are entitled to assert $\ulcorner A \vee \neg A \urcorner$ when and only when A is decidable, in the sense of being either provable or refutable. As we have seen, the Wittgenstein of the *Philosophical Remarks* takes decidability to be the test for whether a mathematical statement has a sense, i.e. qualifies as a proposition. So Brouwer and the Wittgenstein of *Philosophical Remarks* will agree as to which instances of Excluded Middle are assertible. When $\ulcorner A \vee \neg A \urcorner$ is not assertible, though, they will offer different explanations of why not. Brouwer will say it is because A is not guaranteed to have a truth-value. Wittgenstein will say it is because A lacks a sense.

In RFM, Wittgenstein is less explicit than in PR or AWL that he wishes to deal with Brouwer's examples in this way. Implicitly, though, he takes the same line. 'In the law of excluded middle', he writes in §12 of Part V, 'we think we have already got something solid, something that at any rate cannot be called in doubt. Whereas in truth this tautology has just as shaky a sense (if I may put it like that), as the question whether *p* or *~p* is the case' (271).[5] The Wittgenstein of RFM clearly regards the question whether 0123456789 occurs somewhere in the expansion of π as 'shaky'. He deems the question 'queer' (*seltsam*) and says we are led to ask it precisely because we are in the grip of 'the false picture of a completed expansion' of an irrational number (§9, 266, 267).

What the discussion in Part V adds to the earlier doctrine is some explanation of why this question and others like it fail to make sense. Explanation is surely needed here for, at first blush, the question seems to be entirely intelligible. I think we may distinguish two main strands in Wittgenstein's attempt to show that it is not.

5. Section and page references in the rest of this section are to Part V of RFM.

(1) In the first strand, Wittgenstein tries to undermine the most obvious source of confidence that our question makes sense—namely, that we can easily envisage finding ourselves in circumstances where we would return a positive answer to it. We look down a computer print-out of the first one million digits in the expansion of π and—lo and behold—we spot a segment 0123456789. So, to the question 'Does that segment occur somewhere in the expansion of π?', we confidently answer 'yes'. Wittgenstein allows that we would answer the question affirmatively in such a circumstance, but he insists that this does not show that the question possesses a determinate sense:

> If someone says: 'But you surely know what "this pattern occurs in the expansion" means, namely *this*'—and points to a case of occurring,—then I can only reply that what he shows me is capable of illustrating a *variety* of facts. For that reason I can't be said to know what the proposition means just from knowing that he will certainly use it in this case. (§13, 271)

The immediate point here may be Wittgenstein's familiar observation that a single case fails to determine a rule. But his discussion later in Part V of the difference between constructive and non-constructive existence proofs provides more substantial support for the thesis that there are genuinely different interpretations of 'This pattern occurs somewhere in the expansion'.[6] On one interpretation, the only possible ground for asserting the statement would be the identification of the pattern at a specific place in the expansion, as when we spot 0123456789 on the print-out. But there is another interpretation under which the statement also admits of non-constructive proof:

6. I pass over Wittgenstein's suggestion (in §9) that the question is indeterminate in sense because 'the further expansion of an irrational number is a further expansion of mathematics' which calls for 'decisions' about how inherently indeterminate mathematical concepts and rules are to be determined or interpreted. Some mathematical concepts are indeterminate, and as a result some apparently well posed mathematical questions may well lack a determinate sense. For example, it is plausible to maintain that further determination of the concept *set* (or *real number*) is needed before the Generalized Continuum Hypothesis (or the Riemann Hypothesis) qualifies as a well-defined mathematical problem. In these cases, we should agree with Wittgenstein that 'the question…changes its status, when it becomes decidable. For a connection is made then, which formerly *was not there*' (266-7). It is, however, implausible to hold that a conceptual advance of this kind is involved in expanding an irrational number. The rule for writing down the expansion of π is clear and straightforward—a computer may be programmed to follow it—so it is misleading for Wittgenstein to describe this case as one where the 'ground for the decision…has yet to be invented' (*ibid.*).

> A proof that 777 occurs in the expansion of π, without showing where, would have to look at this expansion from a totally new point of view, so that it showed e.g. properties of regions of the expansion about which we only knew that they lay very far out. Only the picture floats before one's mind of having to assume as it were a dark zone of indeterminate length very far on in π, where we can no longer rely on our devices for calculating; and then still further out a zone where in a *different* way we can once more see something. (§27, 284)

The classical mathematician allows non-constructive existence proofs, so he is committed to trying to make sense of the possibility (or apparent possibility) that Wittgenstein sketches in §27. According to Wittgenstein, though, the conditions for making sense of a mathematical proposition are exacting. One needs to 'command a clear view of its application' (§25, 283)—clearly a tall order in the present case. Moreover, the statement in question is liable to engender an *illusion* of understanding. 'This pattern occurs somewhere in the expansion' has the form of an existentially quantified statement, and one is apt to think one understands it because one understands the existential quantifier and understands the relevant matrix instances (in this case, statements of the form 'An instance of the pattern is found starting at the *n*th place'). However, 'the understanding of a mathematical proposition is not guaranteed by its verbal form, as is the case with most non-mathematical propositions', for 'the mathematical general does not stand in the same relation to the mathematical particular as elsewhere the general to the particular' (§25, 282 & 284). At least, this is so in classical mathematics. The classical mathematician allows that someone may prove that a given segment occurs somewhere in an infinite series even when there is no possibility of finding out where. A thinker understands a mathematical proposition to the extent that he knows 'what to do with it', and what one can do with the conclusion of a non-constructive existence proof is very different from what one can do with the conclusion of a constructive proof (§46, 299). These differences are disguised by the fact that the existential quantifier 'somewhere' figures in both 'This pattern occurs somewhere in the expansion' and 'The mug is somewhere in the cupboard'. But this common 'verbal expression … is a mere shadow [which] keeps mum about the important things (*Hauptsache*)' (§25, 282) The logician's use of the symbol '∃' to formalize both of these quantifiers reinforces the illusion of understanding and is a signal illustration of the 'disastrous invasion' of mathematics by logic (§24, 281). The common 'logical notation suppresses the structure' of two very different sorts of statement (§25, 284).[7]

(2) The strand of argument that I have just traced out is designed to shake our confidence that we do understand such statements as '0123456789 occurs somewhere in the expansion of π'. In the second strand, Wittgenstein argues that the claim that we always understand such statements can be maintained only at the price of assimilating the infinite to the finite—the very mistake that both he and Brouwer discern in the higher mathematics of their day. As we have seen, we have a clear apprehension of one sort of ground for asserting our statement—viz., the sort of ground we acquire when we spot 0123456789 in the expansion of π. In §12 of Part V, though, Wittgenstein puts his finger on another reason why this sort of knowledge does not give us the understanding that we seek:

> For how do I know what it means to say: the pattern … occurs in the expansion? Surely by way of examples: which show me what it is like for … [to occur]. But these examples do not show me what it is like for this pattern *not* to occur in the expansion![8]
>
> Might one not say: if I really had a right to say that these examples tell me what it is like for the pattern to occur in the expansion, then they would have to show me what the opposite means. (§12, 271)

This suggests the following argument. In order to attain a clear conception of what it is for P to be the case, one needs to attain a conception of what it is for P *not* to be the case. *Eadem est scientia oppositorum*, as the medieval logicians put it. In the present case, though, we seem to lack the negative side of the story. Or rather, quoting §11 this time,

> To say of an unending series that it does *not* contain a particular pattern makes sense only under quite special conditions.
>
> That is to say: this proposition has been given a sense for certain cases.

7. Cfr. §46 again: 'The curse of the invasion of mathematics by mathematical logic is that now any proposition can be represented in a mathematical symbolism, and this *makes* us feel obliged to understand it. Although of course this method of writing is nothing but the translation of vague ordinary prose' (299).

8. The italicized 'not', although clearly present in Wittgenstein's manuscript, is erroneously omitted from both the German and English editions of RFM. (I am very grateful to Professor Joachim Schulte for pointing this out to me.)

Roughly, for those where it is in the *rule* for this series, not to contain the pattern …(268f.)

What happens if we try to make sense of the hypothesis that 0123456789 appears nowhere in the expansion of π when these 'special conditions' do not obtain? Well, that would mean entertaining the hypothesis that no occurrence of 0123456789 is to be found in the entire expansion, even though such an occurrence is not precluded by the rule for expanding π. And that hypothesis is incoherent on the view of the infinite that Brouwer and Wittgenstein share. It amounts to the absurd hypothesis that the expansion merely *happens* not to contain any instance of 0123456789. In the words of §18:

> Does it make sense to say: 'While there isn't any rule forbidding the occurrence, as a matter of fact the pattern does not occur?'—And if this does not make sense, how can the opposite make sense, namely, that the pattern does occur?
>
> Well, when I say it occurs, a picture of the series from its beginning up to the pattern floats before my mind—but if I say that the pattern does *not* occur, then no such picture is of any use to me, and my supply of pictures gives out…
>
> The queer thing about the alternative 'ϕ occurs in the infinite series or it does not', is that we have to imagine the two possibilities individually, that we look for a distinct idea of each, and that *one* is not adequate for the negative and for the positive cases, as it is elsewhere (277f.).

Thus, when Wittgenstein's 'special conditions' do not obtain, we can attain no clear conception of what is involved in the negative case's being true.

4. *How an intuitionist should reply*

On Wittgenstein's view, then, the conception of the infinite that he and Brouwer share exposes as senseless statements saying that this or that pattern occurs in an infinite decimal expansion, except in the special case when the hypothesis that it does may be proved or refuted. Since Brouwer holds that his intuitionistic logic applies to such statements, he is committed to ascribing a sense to them. How might an intuitionist reply to Wittgenstein's arguments?

What he needs to do is to *attach* a coherent sense to statements of the problematical kind. As Wittgenstein in effect concedes, there is no great difficulty attaching sense to a statement 'ϕ occurs in the infinite series', so long as we understand it in such a way that its ultimate grounds are constructive proofs. So understood, we know in what circumstances we shall be entitled to assert the statement (viz., when we know that ϕ occurs at such-and-such a point in the series) and we also know 'what to do' with such an assertion (viz., look at the proof to discover where ϕ occurs). This method does not extend to attach a sense to our statement, if it is also supposed to admit of a non-constructive proof; but that is not a problem for an intuitionist.

How, though, may we attach sense to 'ϕ does *not* occur in the series'? The key to the intuitionist's answer is his denial that *eadem est scientia* invariably constrains the relation between a statement and its negation. One does not always need a conception of what would be the case if not *P* in order to have a conception of what would be the case if *P*. Rather, one's knowledge of what would be the case if not *P* may draw upon prior knowledge of what would be the case if *P*. So it is in the present case. *Ex hypothesi*, we have a conception of what it would be for ϕ to occur at some identifiable place in the series—identifiable, that is, by means of a mathematical construction. Drawing upon that conception, we can then form the notion of a proof that establishes that no such construction is possible. Such a proof will be the ground for asserting 'ϕ does *not* occur in the series'. Moreover, we know what to do with such an assertion: on its strength, we can set aside for ever any possibility of finding ϕ in the series.

This, in outline, is how the intuitionist should answer the arguments sketched in §3. The answer also shows how to reply to some of Wittgenstein's additional criticisms. Like many critics since, he worries that what the intuitionist refuses to assert is not the 'real' Law of Excluded Middle—i.e., is not Excluded Middle as the classical logician understands it. On the intuitionist's understanding of the statements, 'ϕ occurs in the series' is tantamount to 'It follows from the laws of mathematics that ϕ occurs in the series', and 'ϕ does not occur in the series' is tantamount to 'It follows from the laws of mathematics that ϕ does not occur'. And yet: 'The opposite of "there exists a law that *p*" is not: "there exists a law that *~p*". But if one expresses the first by means of *P*, and the second by means of *~P*, one will get into difficulties' (§13, 272). Or again: 'If "you do it" means: you must do it, and "you do not do it" means: you must not do it—then "Either you do it, or you do not" is not the law of excluded middle (§17, 275). It is certainly not the Law of Excluded Middle as the classical logician understands it, but that cannot be a legitimate criticism. Wittgenstein agrees with Brouwer that any attempt to apply classical negation to 'ϕ occurs in the series' will result in nonsense. So the intuitionist cannot be faulted for trying to articulate a non-classical conception of negation, which in turns yields a non-classical reading of the Law of Excluded Middle. On that conception, $\ulcorner \neg A \urcorner$ is inherently a more complex statement than A, so it should be no surprise that $\ulcorner \neg\neg A \urcorner$ does not always entail A, or that $\ulcorner A \vee \neg A \urcorner$ is not always assertible.

Our analysis also brings out the depth of the gulf that separates Brouwer's case for intuitionism from Dummett's. On Brouwer's view, we are driven to interpret mathematical statements in terms of constructions because the attempt to apply a classical interpretation, which respects *eadem est scientia*, leads ineluctably to an incoherent view of the infinite. His case, then, is specific to higher mathematics. It is not, and cannot be, the harbinger of a general argument in favour of casting semantic theories in terms of assertibility-conditions rather than truth-conditions.

5. *A lasting legacy of the* Tractatus

At the heart of the dispute between Brouwer and Wittgenstein lies a disagreement about the conditions that a form of words must satisfy in order to qualify as a proposition—that is, to be an intelligible statement to which the laws of logic apply. The following formulation of the disagreement may be helpful. Let us assume that *denying* a proposition is logically equivalent to asserting its negation: both classical and intuitionist logicians will grant this assumption. Let us then say that a statement has a *back* when an assertion of it *ipso facto* amounts to a denial of some other statement. Both classical and intuitionist logicians assume that any statement *has* a negation. A statement with a back will also *be* a negation, or be equivalent to one. That is, A has a back if and only if, for some statement B, A is equivalent to $\ulcorner \neg B \urcorner$; to assert A will be to deny B. The locus of dispute between Brouwer and Wittgenstein is then the following thesis:

(B) Every proposition has a back, i.e., every proposition is the negation of some other proposition.

Like any intuitionist, Brouwer cannot assert (B): were he to assert it, intuitionistic propositional logic would collapse into classical propositional logic. The reason is this. For any formula B, the triple negation $\ulcorner \neg\neg\neg B \urcorner$

is intuitionistically equivalent to the single negation $\ulcorner\neg B\urcorner$. Suppose, then, that A has a back. Then, for some B, A is equivalent to $\ulcorner\neg B\urcorner$, so that $\ulcorner\neg\neg A\urcorner$ is equivalent to $\ulcorner\neg\neg\neg B\urcorner$. By the result about triple negations, this means that, whenever A has a back, A is intuitionistically equivalent to $\ulcorner\neg\neg A\urcorner$. So, if an intuitionist were to assert (*B*), he would be committed to taking each proposition to be equivalent to its own double negation. That would suffice to collapse intuitionistic propositional logic into classical logic.

Wittgenstein, by contrast, had a long-standing and deep-seated commitment to (*B*). When we understand a proposition, he wrote in the 'Notes on Logic' of September 1913, 'we know what is the case if it is true and what is the case if it is false' (NB, 94). In this way, any proposition is associated with true and false 'poles'. To accept the true pole is *ipso facto* to reject the false pole. The negation operator, on Wittgenstein's account, simply reverses the poles, so asserting that *P* is *ipso facto* denying that not *P*, just as (*B*) has it. *Eadem est scientia* follows. This is why the *Tractatus* makes no room for doubting the equivalence between a proposition and its double negation. These say the same thing (TLP 5.44); indeed, in a fully perspicuous symbolism, double negations would vanish (TLP 5.254). In any event, the universal equivalence of a proposition and its double negation suffices (given very weak assumptions about the logic of disjunction) to ensure the validity of every instance of Excluded Middle.[9]

But is it really a universal requirement that any fully intelligible statement should have a back? (*B*) has great initial plausibility: it is at first hard to see how a statement could have a determinate content unless it is a determinate matter what it excludes. And our reluctance to deviate from (*B*) explains, I think, why so many reasoners are willing to apply classical logic even to statements whose bivalence they find doubtful. In the previous section, though, we saw reasons why statements involving infinity might be exceptions to (*B*). In asserting '0123456789 occurs somewhere in the expansion of π', there is nothing that one is thereby denying. In particular, one is not thereby denying '0123456789 occurs nowhere in the expansion'. In order to understand a negated statement one must understand its negand, but not necessarily *vice versa*.

In the light of that, I do not think that anyone could claim that (*B*) is *obviously* correct. So Wittgenstein has not shown that any attempt to attach sense to statements of the problematical kind must fail. To say as

9. Let A be any statement. Since $\neg(A \vee \neg A)$ intuitionistically entails a contradiction (see *n*.2), $\neg\neg\,(A \vee \neg A)$ is a theorem of the intuitionistic propositional calculus. So if all double negations were eliminable, we would always have $A \vee \neg A$.

much, of course, is not to say that Brouwer or any other intuitionist has succeeded in attaching sense to those statements. To show that, one would need to elaborate the putative sense to the point where it clearly provides a coherent alternative to the classical account that Brouwer and Wittgenstein both reject. Like them, I regard that classical account as deeply suspect. So I regard the open question here—whether the intuitionist can succeed in articulating an alternative sense, or whether we must follow Wittgenstein in deeming such statements to be senseless—as one of the most important in the philosophy of mathematics.[10]

References

Brouwer, Luitzen 1923: "Über die Bedeutung des Satzes vom ausgeschlossenen Dritten in der Mathematik, insbesondere in der Funktionentheorie". *Journal für die reine und angewandte Mathematik* 154, 1–7. Page references are to the translation by Stefan Bauer-Mengelberg and Jean van Heijenoort in: van Heijenoort (ed.), 1967, 335-41.

— 1928: "Mathematik, Wissenschaft und Sprache". *Monatshefte für Mathematik und Physik* 36, 153–64. Page references are to the translation by William Ewald in: Ewald (ed.), 1996, 1170–85.

Ewald, William (ed.) 1996: *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, volume II. Oxford: Clarendon Press.

Gödel, Kurt 1931: "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I". *Monatshefte für Mathematik und Physik* 38, 173–98. Page references are to the translation by Jean van Heijenoort in: van Heijenoort (ed.), 1967, 596–616.

Heijenoort, Jean van 1967: *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*. Cambridge, MA: Harvard University Press.

Kolmogorov, Andrey 1925: "О принципе tertium non datur". *Мамеламическуў Сборнук* 32, 646–67. Page references are to the translation by Jean van Heijenoort in: van Heijenoort (ed.), 1967, 416–37.

Pitcher, George 1964: *The Philosophy of Wittgenstein*. Englewood Cliffs, NJ: Prentice-Hall.

---

10. I am much indebted to Lucy Baines, Hanjo Glock, Joachim Schulte, and Göran Sundholm for their comments on drafts of this paper.

Special Topic II

QUINE

Grazer Philosophische Studien
89 (2014), 111–112.

INTRODUCTION

Quine visited Brazil for the first time in 1942. He lectured on logic, in Portuguese, as a Visiting Professor at the University of São Paulo. Seventy years later, in 2012, his deep impact on Analytic Philosophy in Brazil was commemorated by a symposium at the Universidade Federal Fluminense in Niterói, Rio de Janeiro. The papers collected here were originally read at this symposium. They are dedicated both to the interpretation and the critical assessment of Quine's philosophy.

The first three papers focus on Quine's philosophy of language. In "Significance in Quine", Peter Hylton aims to clarify Quine's views about meaningfulness, especially his position with regard to the verificationist conception of meaning and meaningfulness. Rogério Severo's paper, "Are There Empirical Cases of Indeterminacy of Translation?", defends Quine's thesis of the indeterminacy of translation against the common objection that it has never been backed up by any positive empirical evidence. In "Some Critical Remarks on Quine's Thought Experiment of Radical Translation", Oswaldo Chateaubriand points out to some theoretical and empirical weaknesses of Quine's argument for the indeterminacy of translation considered as a thought experiment.

The remaining three papers are mainly on Quine's ontology. In "A Tension in Quine's Naturalistic Ontology of Semantics", Dirk Greimann argues that Quine's naturalistic limitation of the facts of semantics to the facts about verbal behavior conflicts with his principle of ontological commitment, broadly construed. Guido Imaguire's paper, "In Defense of Quine's Ostrich Nominalism", aims to show that Quine's nominalist conception of the predicates is the best answer to the Platonic One over Many argument. Finally, in "Quinean Worlds: Possibilist Ontology in an Extensionalist Framework", Pedro Santos constructs an ontology for modal logic that is inspired by Quine's remarks in "Propositional Objects" and that respects his ontological constraints, specially the principle of extensionalism and the principle of individuation.

I am grateful to all the contributors to this volume for their cooperation and encouragement. Special thanks are due to Peter Hylton, the keynote speaker of the conference. I am obliged to the Universidade Federal

Fluminense for financial support. Last but not least, I would also like to thank the editors of *Grazer Philosophische Studien* for making this volume possible.

December 2013
Dirk GREIMANN

Grazer Philosophische Studien
89 (2014), 113–133.

SIGNIFICANCE IN QUINE

Peter HYLTON
University of Illinois at Chicago

Summary

This essay concerns Quine's views about significance. The first section discusses 'Meaning in Linguistics', and argues that the strategy of that essay cannot provide the basis of a notion of meaningfulness which has any philosophical role to play. The next two sections argue that, in spite of some evidence to the contrary, Quine is not committed to any version of verificationism. The final section takes up the question of the scientific language which is the subject of much of Quine's discussion. I argue that he is not concerned to impose on that language any controversial distinction between the meaningful and the meaningless.

I.

My subject, to be clear, is not the significance *of* Quine but rather the *idea* of significance as it figures *in* Quine's work. Significance, as I am thinking of it here, is an attribute of linguistic expressions: an expression's being significant, or meaningful, is of course the opposite of its being meaningless, or nonsense. The idea of nonsense plays an extremely interesting role in the history of analytic philosophy, from Frege on. Our focus is on Quine, but it will be useful to begin with a very brief discussion of Carnap's views.

The view that metaphysics is not merely useless, but nonsense, meaninglessness in a clear-cut sense, is one of two crucial ideas that Carnap attributes to the influence of Wittgenstein. Carnap's use of the idea of nonsense is seen most clearly in *Scheinprobleme in der Philosophie* (Carnap 1928/1967), which was written in 1927, when Wittgenstein's influence on Carnap was at its height. In that book the idea is based on a notion of verifiability. (Carnap later speaks of 'Wittgenstein's principle of verifiability'; Carnap 1963, 45) According to *Scheinprobleme,* a statement has 'factual content' if 'experiences which would support [either that statement or its negation] are at least conceivable'; expressions which lack factual content 'must under no circumstances be considered meaningful' (Carnap 1928/1967, 327f.). Carnap's later work qualifies this view in various ways, and comes to put more emphasis on violations of logical syntax and less on verifiability, but he continues to speak of 'pseudo statements' (*Schein-Aussagen*), 'pseudo problems' (*Schein-Probleme*), and so on. In an essay published in1950, Carnap says of philosophers who ask ontological questions: 'Unless and until they supply a clear cognitive interpretation, we are justified in our suspicion that their question is a pseudo-question …' (Carnap 1950/1956, 209).

The other crucial idea that Carnap attributes to the influence of Wittgenstein is, of course, that the truths of logic, including mathematics, are tautologous or, in the terminology that he came to favour, analytic (see Carnap 1934/1937, 44). In his essay, 'Carnap and Logical Truth', Quine links these two ideas. In sketching the views of Carnap, and other members of the Vienna Circle, he says: 'Metaphysics was meaningless through misuse of language; logic was certain through tautologous use of language' (Quine 1963/1976, 108). In spite of this explicit linkage, however, the two issues play very different roles in Quine's work. The second, of course, comes in for a good deal of discussion, chiefly under the heading of 'analyticity' or (more or less interdefinably) 'synonymy'. Quine's discussion of analyticity and synonymy is the pivotal point of break with Carnap, which leads to his articulation of his own rival views. The issue of meaninglessness, or significance, receives much less attention in Quine's work, and the various remarks he makes may seem to point in different directions.

In 'Two Dogmas of Empiricism' (Quine 1953a), Quine seems to repudiate verificationism on the grounds of holism. 'The Problem of Meaning in Linguistics' (Quine 1953b), written around the same time as 'Two Dogmas of Empiricism', does not discuss verificationism but does put forward a favourable view of the general idea of meaningfulness, or significance, explicitly contrasting it with analyticity. In later works, moreover, there are several passages in which he seems to endorse verificationism, presumably in a version compatible with holism. In two places, indeed, he puts forward what seems to be an argument for that position; in one of those passages he also seems to employ a verificationist criterion of meaningfulness to dismiss a rival position. Not surprisingly, in view of these discussions, a number of notable commentators have attributed a verificationist view to him: Roger Gibson, for example, says, in the context of a discussion of cognitive meaninglessness, 'Quine retains the positivists' criterion of verifiability' (Gibson 1988, 21). Michael Dummett is equally direct, saying: 'Quine's model of language is as verificationist as the positivist model'.[1]

My aim, in the present essay, is, to come to a clear understanding of just what Quine's views are about significance, or its opposite, nonsense—to get clear on the fate that the idea of nonsense meets at Quine's hands. The first section will discuss 'Meaning in Linguistics', and argue that the strategy of that essay cannot provide the basis of a notion of meaningfulness which has any philosophical role to play (and, arguably, cannot provide the basis for any notion of meaningfulness at all). The next two sections will discuss Quine's apparent commitment to what we might call holistic verificationism. Section II will present some evidence that Quine holds that view and that he puts forward an argument for it, based on how language is acquired. Section III will seek to show that this argument for it is, by Quine's own standards, not a good one, and that his other views do not commit him to any form of verificationism—although he may indeed have been tempted by it at times. The discussion to this point concerns natural language. The final section takes up the question of the scientific language which is the subject of much of Quine's discussion. Some idioms are excluded from that language, so sentences which use them are, of course, not meaningful sentences of that language. Within the grammatical sentences of the language, however, Quine does not wish to distinguish the meaningful from the meaningless; in particular, he explicitly rejects the idea that it would be desirable to impose a verificationist criterion of meaningfulness on the language. My conclusion is therefore that, in spite of some signs of his having been tempted by verificationism, the idea of nonsense has no significant role in Quine's mature thought.

II.

In 'The Problem of Meaning in Linguistics', Quine is at some pains to separate the issue of significance, on the one hand, from the issue of synonymy and analyticity on the other. The former can be phrased in terms

1. Dummett 1978, 379. For this reference I am indebted to Panu Raatikainen (Raatikainen 2003). The conclusions of this latter essay are in line with my own, although the argument differs significantly.

of an expression's *having a meaning*; the latter in terms of two expressions' *having the same meaning*, or *being alike in meaning*; so it may seem as if there is a single issue here, or two issues that are closely connected, as different aspects of a single question of meaning. Quine denies this, however. He recommends that we avoid the term 'meaning' in each case. In approaching the issue of synonymy, he recommends that we 'treat this whole context [being alike in meaning] in the spirit of a single word "synonymous", thus not being tempted to seek meanings as intermediate entities' (Quine 1953b, 48). He recommends a similar approach to the idea of a linguistic form's having a meaning: 'treat the context "having a meaning" in the spirit of a single word, "significant", and continue to turn our backs on the suppositious entities called meanings' (*loc. cit.*). Similarly, he argues in *Word and Object* against what he calls 'the fallacy of subtraction' which he describes as follows: 'it is argued that if we can speak of a sentence as meaningful then there must be a meaning that it has, and this meaning will be identical with or distinct from the meaning that another sentence has' (Quine 1960, 206).

Quine thus wishes to block any inference from the claim that meaningfulness makes clear sense to the claim that synonymy (and with it analyticity) also makes clear sense. The reason for this is that he accepts the first claim but is dubious, to say the least, about the second. In 'Two Dogmas of Empiricism', the second chapter of *From a Logical Point of View*, he argues, of course, against the acceptability of the idea of analyticity and, with it, the idea of synonymy; in particular, he argues that those ideas are in need of clear behavioural criteria, and that the philosophers who employ them have no such criteria. In 'The Problem of Meaning in Linguistics', the third chapter of *From a Logical Point of View*, by contrast, he sets out to investigate whether there are such behavioural criteria for significance, comes up with an affirmative answer, and sketches the relevant criterion. On that basis, Quine concludes that one aspect of 'the problem of meaning', namely 'the aspect of having a meaning' is in 'halfway tolerable shape' (Quine 1953b, 56). It is explicit that in saying this he means to draw a contrast between significance, on the one hand, and synonymy, on the other (Quine 1953b, 64).

A closer examination of 'The Problem of Meaning in Linguistics', however, suggests that Quine's acceptance of the idea of meaningfulness there may not amount to very much. The significant expressions of a language, Quine says, are those which are actually uttered by its speakers together with 'those which *could* be uttered without reactions suggesting bizarreness of idiom' (Quine 1953b, 53). The idea of a reaction 'suggesting bizarreness of idiom' is one that Quine here seems to find acceptable on the face of it, without further discussion. The point he finds potentially problematic is the 'could'. This, he says, can be explained in terms of the grammarian's attempt to systematize actual utterances of the language in the grammatically simplest way. That attempt will issue in a grammar which includes, as part of the language, indefinitely many sequences of expressions which have not in fact been uttered; such sequences will count as meaningful expressions of the language (pending tests which might lead to revision of the proposed grammar). Quine sees this as a point about scientific method quite generally, not only in linguistics: 'Our basis for saying what "could" be generally consists, I suggest, in what *is* plus *simplicity* of the laws whereby we describe and extrapolate what is. I see no more objective way of construing the *conditio irrealis*' (Quine 1953b, 54). Claims about what could be uttered without provoking a bizarreness reaction are also testable in the obvious way. So Quine accepts that we have at least an approximate behavioural criterion for significance.

Various criticisms could be made of this criterion, from a more or less Quinean point of view. Most obviously, it seems quite problematic to assume that we can distinguish reactions which suggest 'bizarreness of idiom' from those which indicate some other form of puzzlement over an utterance, e.g. that it is obviously false or irrelevant to the topic in hand. The ordinary utterance of terms such as 'Nonsense!' or 'Absurd!' makes no such distinction. Speakers frequently use such labels for views with which they strongly disagree, or think are foolish, without meaning to imply that the sentences they reject are simply gibberish like 'Boo goo foo'. So one may wonder whether there really is a behavioural notion which will play the role that Quine needs here; and indeed Quine himself expresses doubt of just this sort in later work.[2]

We may, however, pass over this issue, and various other possible criticisms, for there is a more fundamental point. Even if fully successful on its own terms, the position that Quine puts forward in 'Meaning in Linguistics' would not yield a notion of meaningfulness which would be

2. In 'Philosophical Progress in Language Theory' he says 'when we try to infer meaninglessness from verbal behavior, we cannot easily distinguish it from mere extravagance' (Quine 1970b, p. 7). In *Philosophy of Logic* he says '[t]o speak of "what could occur in normal speech" is... objectionably vague. The vagueness … lies less in the 'could' than in the 'normal', or in their combination' (Quine 1970a, p. 21).

of any philosophical interest or use. The crucial point is that the behavioural criterion that Quine suggests will not give us a basis on which to dismiss as meaningless anything which is uttered by native speakers with serious intent, or which could be so uttered. Thus we might hope, for example, to use a criterion of meaningfulness, in old Viennese fashion, to eliminate speculative metaphysics or theology. But the sentences of those subjects have been uttered, presumably with serious intent, by competent speakers of their respective languages; hence the behavioural criterion gives us no basis on which to say that those sentences are meaningless. True, the uninitiated, confronted with a sentence taken from a metaphysical or theological discussion, may well react in ways suggesting they find the utterance bizarre. But the same may hold, surely, of an uninitiated person confronted with a theoretical sentence of physics or mathematics. In each kind of case, however, novices can be educated. Not only in the case of mathematics and physics, but also in the case of metaphysics and theology, some can be initiated, step-by-step, into the practice of uttering such sentences, and responding to them when others utter them, in ways that the already initiated accept. So there is in Quine's notion of significance no basis on which to dismiss such subjects as meaningless.

The point here is a general one: a criterion of linguistic significance based purely on the linguistic habits and dispositions of users of the language cannot provide a basis on which to dismiss any utterance which those users (or a significant number of them) are actually disposed to make, or to hear without manifesting a bizarreness reaction. A criterion of that sort can do nothing to separate the utterances which users make into the significant and the meaningless; any utterance that is seriously made is significant by this criterion, just in virtue of that fact. So Quine's behavioural criterion of significance will not function as a philosophical tool or weapon. At most, it seems, it would yield a notion comparable to the notion of analyticity that Quine ends up accepting[3]: unobjectionable, but not suited to any serious philosophical purpose.

3. See, for example, 'Two Dogmas in Retrospect', where Quine says 'I recognize the notion of analyticity in its obvious and useful but epistemologically insignificant applications' (Quine 1991, 271).

III.

If Quine holds a notion of meaningfulness that is to be at all philosophically useful, it must be on some other basis than that which he puts forward in 'Meaning in Linguistics'. In several passages Quine seems to endorse such an alternative basis, namely some version of verificationism.[4] One such passage occurs in his 'Reply to Roger F. Gibson Jr.' (Quine 1986). Gibson had discussed an idea, which he attributes to Dagfinn Føllesdal, that 'the indeterminacy of translation follows from holism and verification theory of meaning' (Quine 1986, 155). Both Gibson and Føllesdal have doubts about this argument because they have doubts about the verification theory of meaning. Quine, however, takes a different view:

> I find [this argument for indeterminacy] attractive. The statement of verificationism relevant to this purpose is that "evidence for the truth of a sentence is identical with the meaning of the sentence"; I submit that if sentences in general had meanings, their meanings would be just that. It is only holism itself that tells us that in general they do not. (Quine 1986, 155f.)

It is unclear exactly what we should make of a statement from Quine starting 'if sentences in general had meanings...', but it seems to indicate, at least, a considerable degree of sympathy for the general idea behind verificationism.

Perhaps the strongest evidence for ascribing a verificationist criterion of meaningfulness to Quine is a passage in 'Epistemology Naturalized' (Quine 1969).[5] Quine seems, at least, to use that criterion to eliminate a view opposed to his own and to defend this step by arguing for verificationism. The context is a discussion of the indeterminacy of translation. Quine claims that it is highly likely that there could be different translations which would 'deliver the same empirical implications for the theory as a

4. Among the Logical Empiricists it was controversial exactly how the relevant criterion should be formulated—whether in terms of conclusive verifiability, or of confirmability, and so on. See Hempel, 1965. I shall not be concerned with this issue: for present purposes we may assume that what is at stake is a relatively weak criterion formulated in terms of confirmability.

5. There is also a passage from the end of 'Limits of Knowledge' in which Quine also seems to endorse a verificationist position: 'The relation of language to observation is often very devious, but observation is finally all there is for language to be anchored to. If a question could in principle never be answered, then, one feels, that language has gone wrong; language has parted its moorings, and the question has no meaning' (Quine 1966/1976, 67). It is difficult to know how much weight to put on this, both because the endorsement is less than explicit ('one feels that') and because the essay is semi-popular in nature, having originally been written for broadcast as a talk on Radio Canada.

whole' (Quine 1969, 80). In that case, he says, 'there can be no grounds for saying which of two glaringly unlike translations of individual sentences is correct'. Quine holds that we should conclude from this situation that indeterminacy of translation holds. He acknowledges, however, that '[f] or an uncritical mentalist, no such indeterminacy threatens' (*loc. cit.*). It is at this point that he seems to deploy verificationism to argue against this mentalist view: 'When on the other hand we take a verificationist theory of meaning seriously, the indeterminacy would seem to be inescapable' (*loc. cit.*).

Recognizing that some readers will oppose indeterminacy, he asks: 'Should the unwelcomeness of the conclusion persuade us to abandon the verification theory of meaning?' (81). His answer is forthright: 'Certainly not.' (*loc. cit.*). He proceeds to put forward what is, in context, clearly intended as an argument for the verification theory:

> The sort of meaning that is basic... to the knowledge of one's own language, is necessarily empirical meaning and nothing more. A child learns his first words and sentences by hearing and using them in the presence of appropriate stimuli. These must be external stimuli, for they must act both on the child and on the speaker from whom he is learning. Language is socially inculcated and controlled; the inculcation and control turn strictly on the keying of sentences to shared stimulation … Surely one has no choice but to be an empiricist so far as one's theory of linguistic meaning is concerned. (*loc. cit.*)

On this account, the 'keying of sentences to external stimuli' is the essential thing. Learning a language is learning to link utterances to external stimuli in the same sorts of ways as do those from whom one is learning it. The requirement of verifiability presumably follows because it is only insofar as an utterance is thus linked to external stimuli that it counts as a piece of meaningful language at all, and the linkage will guarantee that the utterance is in some way answerable to external stimuli.

The connection between language and external stimuli is central not only to the learning of language but also to the way in which utterances in the language thus learned are answerable to evidence. On Quine's account, the two issues thus become one. As he puts it towards the end of 'Epistemology Naturalized': '...epistemology now becomes semantics. For epistemology remains centered as always on evidence, and meaning remains as always centered on verification; and evidence is verification' (Quine 1969, 89).

In *Roots of Reference* (Quine, 1974), Quine argues, in similar vein, that the evidential relation is identical to what he calls 'the semantical relation' (p. 38). We can understand the relation of sentences to the observations which confirm them or disconfirm them by examining how language is learned: 'we learn the language by relating its terms to the observations that elicit them.... By exploring [this learning process], science can in effect explore the evidential relation between science itself and its supporting observations' (Quine 1974, 37). Again, he suggests that this supports verificationism:

> The two roles of observations, their role in the support of theory and their role in the learning of language, are inseparable. Observations are relevant as evidence for the support of theory because of those very associations, between observable events and theoretical vocabulary, whereby we learn the theoretical vocabulary in the first place. Hence, of course, the commonplaces of the verification theory of meaning. The meaning of a sentence lies in the observations that would support or refute it. To learn a language is to learn the meaning of its sentences, and hence to learn what observations would count as evidence for and against them. The evidence relation and the semantical relation of observation to theory are coextensive. (Quine 1974, 38)

More explicitly than in 'Epistemology Naturalized', Quine in *Roots of Reference* qualifies this endorsement of verificationism with holism. He emphasizes that most sentences do not relate to observations individually, but only via other sentences, so '[t]he evidence relation is thus intricate and indirect' (*loc. cit.*). But, he claims, the semantical relation is equally intricate and indirect 'since we learn the language only partly by associating terms or sentences directly with observation, and partly by linking them to one another' (*loc. cit.*). Hence, he concludes, '[t]he evidence relation, in all its intricacy, and the semantical relation, in all its intricacy, are coextensive still' (*loc. cit.*).

All of this may suggest that Quine's position on meaningfulness is, so to speak, holistic verificationism. We learn language from other people, in observable circumstances. So meaningful language must be related to observations which constitute both the basis on which it can be learned and the evidence for and against the sentences thereby learned. So some form of verificationism must be correct. But it is *holistic* verificationism because in many cases the relation of a given sentence to observation will be mediated by other sentences, whose relation to observation may in turn be mediated by yet other sentences, and so on without evident limit—so that it may be difficult, to the point of being impossible, to say in advance what counts as the evidence for a given statement. Nevertheless, according to this view, the way in which language is learned shows that our sentences must stand in some relation to observations which would count as evidence for it or against them, however intricate and difficult to tease out that relation may be.

IV.

The argument set out in the previous section purports to show, on the basis of the way that language is acquired, that verificationism, albeit qualified by holism, must be correct. In particular, it purports to show that only language obeying verificationist constraints could be acquired in the way in which language in general is in fact acquired (according to Quine). This argument, however, seems to me quite unconvincing, as this section will attempt to show. In particular, I shall argue that Quine's sketch of how language might be acquired does not in fact support the idea that the sentences which we come to be disposed to utter must be related to observation, not even if we take account of holism by including indirect relations to observation.

One step which Quine discusses in language acquisition is what he calls 'attributive composition'. He uses as an example the sentence 'Mama is smiling' (Quine 1974, 61). Quine assumes that we have learned the two constituent terms as observation sentences. Then, he says, we learn the sentence as a whole 'thanks to the intersecting of the pertinent saliences; the smiling occurs on Mama' (*loc. cit.*). On this account, we learn two terms observationally, and then pick up a mode of combining them, also in more or less observational fashion. Once our infant language learner has made it that far, however, nothing prevents her from going further. Her linguistic creativity may take over; she may say of a pet dog, say, or of a favourite doll, or even of a rock, that it is smiling.[6] These last sentences may be unlikely to meet with the same sort of reinforcing approval as 'Mama is smiling', uttered under appropriate circumstances. Still, the way in which language is acquired does not rule them out, or rule out any of the countless sentences—or apparent sentences—of the same sort.

6. Calvin Trillin reports his then four year-old daughter Abigail as uttering the sentence 'My tongue is smiling.' upon first eating chocolate ice cream. See Trillin 1978/2007, 7.

The point here is a general one. Perhaps the learning of each part of a sentence, and of the modes of combination of such parts in the sentence, is tied, directly or indirectly, to observations. (The more latitude we give to 'indirectly', the more plausible this becomes.) But still the sentence itself may be remote from observation, or even observationally altogether inaccessible—unverifiable, in a word. This kind of thing is not rare or esoteric. Attributive composition by itself allows for any predicate to be ascribed to any noun and it is, as Quine says, 'one of many dyadic constructions on terms' (Quine 1974, 61). Another example he gives is the 'in' construction, as in the sentence 'Mama is in the garden'. If one focuses on the kind of case closest to observation—a small garden, under current observation and without trees, bushes, or huts which might conceal a person—then it may seem as if the 'in' construction will not take us far from observation. But the 'in' construction, once mastered, leads to other sorts of cases, and to sentences which have no evident relation at all to current observation. 'The cat is in the garden' will in some cases be answerable to current observation; in other cases, however, such being the way of cats, it may be wholly elusive to observation. A further step is taken when the child masters ways of talking about other places and other times; sentences remote from current observation then become commonplace. A child who has mastered 'It's raining' and 'Now we're at Grandma's house' may, while at home, utter 'It's raining at Grandma's house' even though nothing in current observations speaks for or against that meteorological report. A sentence such as 'The cat was in the garden last night' may outrun all conceivable future observations. And a child who has been told that heaven is a happy place where good people go when they die may be comforted by being told 'Grandpa is in heaven', or may utter that sentence herself, despite its apparent lack of connection with any potential observations, past, present, or future—its apparent unverifiability.

In going from the observation sentences to even fairly elementary non-observation sentences, the child may thus proceed from sentences subject to the most direct and immediate verification to others whose status is, in this respect, dubious at best. Quine's account of language acquisition is not intended to bridge this sort of gap, nor could it do so. It would, indeed, be more accurate to say that it is intended to show how a gap could open up here. Quite generally, Quine does not depict the child as acquiring language by a rational process, in which confirmation is transmitted from the observation sentences which form the starting point to the theoretical sentences which are eventually mastered. The learner's progress in acquiring language, he says, 'is not a continuous derivation, which, if followed backwards, would enable us to reduce scientific theory to sheer observation. It is a progress by short leaps of analogy.' (Quine 1975, 77f.). What the shortness of the leaps comes to here is that it must be plausible that a child, having acquired one idiom, could succeed in acquiring another; it does not imply or even suggest that verifiability or confirmability by experience will be passed down from one idiom to another acquired on the basis of the first.

In some cases, the process is not merely not rational; it has something disreputable about it. Quine considers how the child is able to acquire the ability to form 'eternal predications', subject-predicate sentences true or false once for all. He takes as an example the sentence 'Snow is white', and says that the child is able to learn it because her response to the word 'snow' is at least to some extent the same as her response to the substance, snow. It is, Quine says, 'the essence of the notorious confusion of sign and object, or of use and mention' (Quine 1974, 68). 'Language', he says, 'is rooted in what a good scientific language eschews … Language is conceived in sin and science is its redemption' (*loc. cit.*). In a similar vein, Quine says that we may learn idioms in contexts where they are eliminable in favour of those already known, and then proceed to use them, by analogy rather than by definition, in contexts from which they are not eliminable (see, for example, Quine 1974, 93ff.). Even if sentences formed by the first kind of use of the idiom are verifiable, there is no reason at all to think that this trait will be inherited by sentences formed by the second kind of use of the idiom.

Mastering an observation sentences, on Quine's account, consists in acquiring a disposition to respond one way or another, depending on one's current neural intake. Thus far, language is as firmly based on observations as anyone could wish; a sign of this is that utterances of observation sentences will meet near-unanimous agreement from others in the same perceptual situation. From that point on, however, the links between language and experience grow looser and the room for disagreement grows greater. We learn language from others on the basis of their overt behaviour in observable situations, yes, but the way in which learning proceeds—with leaps and analogies, and use-mention confusions, and who knows what else?—leaves room for sentences which wholly elude attempts to bring evidence to bear on them, and thus leaves room for irresoluble disagreements. What we learn, out beyond the observation sentences, are not individual sentences but constructions, ways of forming sentences from previously learned components. Being the sort of creatures that we are, we will rapidly be led by these means to utter and to respond to sentences that outstrip our capacity to bring the evidence to bear on them. All of this does not conflict with Quine's schematic account of how language might be acquired but is, to the contrary, implicit in that account.

At this point it may be tempting to ask: does the child who utters such sentences *understand* them? Indeed does the adult who utters an unverifiable sentence understand it? For someone who advocates a clear-cut notion of meaningfulness it is an appropriate question, for understanding is directly correlative with meaningfulness: if a person understands a sentence then that sentence is meaningful for her, and *vice versa*. Quine, however, would reject the question. Once we have fully described the child's abilities to use and to respond to the sentence, there is no further question as to whether those abilities amount to *understanding* the sentence involved. There is no Quinean 'theory of understanding'; 'understanding' is not a sufficiently clear term to support a theory. The closest we can come is a theory of how language is used, and of how it comes to be used in that way—in particular, an account of how language is acquired. That account will not support the idea of an all-or-nothing notion of understanding; the various abilities that the child (or, indeed, the adult) comes to possess with regard to various sentences will, to use a Quinean phrase, grade off. There will be no place to draw a line and say: on this side of the line there is meaningful language, on that side there is meaningless noise. So Quine's account of how language might be acquired does not support any version of a verificationist criterion of meaningfulness—although Quine sometimes seems to suggest that it does.

Even if we first utter a sentence without being able to bring evidence to bear on it, we may, in time, learn to relate it to evidence—as our imagined child will in time learn what counts as evidence for or against the claim that it is raining at Grandma's house. There is, however, every reason to think that our linguistic creativity will continue to outstrip this process, and that some of the sentences we utter, and count ourselves as understanding, may never acquire a firm relation to potential evidence, for or against. The vagaries of language acquisition also leave room for disagreement from one person to another, even if their perceptual experience is not relevantly different. The experience on the basis of which the utterer proceeds does not guarantee that the rest of us will go along, and hence does not guarantee that we will count the utterance as true.

These phenomena are not the same as the holism that Quine counts as telling against sentence-by-sentence verificationism; they are additional. Consider again the story of the child who, more or less out of the blue, utters 'It's raining at Grandma's house'. That child is not responding to experience which is, in some intricate and indirect fashion, evidence for the sentence she utters; she is, rather, simply exercising her linguistic creativity. (Something that Quine's account of language-acquisition leaves room for, and indeed relies upon.) On the other hand, the fact that we speak a holistic language means that some sentences may have a useful role in our knowledge even though their relation to evidence is by no means clear at first sight (and, for the language-learner, may not exist at all when the sentence is first acquired). Hence we have reason to tolerate such sentences, to respond to them more or less as we do to more obviously sensible uses of language, and perhaps to venture to assert them ourselves. Many such sentences do indeed have, or may come to acquire, indirect relations to evidence which make them useful. But the way is left open for other sentences which do not—unverifiable sentences.

I have argued that the argument for verificationism which we seem to find in 'Epistemology Naturalized' and in *Roots of Reference* is quite unconvincing. This fact might prompt us to consider a different interpretation of the passages. Perhaps Quine is there not arguing for a verificationist criterion of meaningfulness but rather for what he elsewhere calls a behaviourist approach to meaning (Quine 1990a, pp. 37f.), which might better be called simply an empiricist approach. Perhaps the passage we discussed from 'Epistemology Naturalized' is not a rejection of 'uncritical mentalism' as unverifiable and therefore meaningless; perhaps Quine is there putting forward what he there calls a 'verificationist' view of meaning, but which might equally be called simply an empiricist view, as a superior alternative to mentalism. On the other hand, his choice of the word 'verificationist' is presumably not capricious, and he would have been well aware that most people reading that word would associate it with the idea of a criterion of meaningfulness. The issue here is whether Quine was, at least for a time, tempted to think that some sort of verificationist notion of meaningfulness followed from an empiricist approach to meaning. For present purposes we can leave this interpretive issue unsettled. As already indicated, I do not think that any such implication in fact holds; nor do I think that it represents his settled opinion on the matter.

V.

To this point we have been focused on natural language and its acquisition. Natural language, however, is not the only kind of language with which Quine is concerned. An important role in his thought is also played by the idea of a regimented language designed to accommodate our knowledge in the clearest and most objective fashion. Nothing we have said so far addresses the role of the idea of meaninglessness in that kind of scientific language.

Since the scientific language is conceived of as a deliberate creation—designed, not evolved—we, its creators, make decisions about what terms it includes. The decisions that Quine advocates would be very restrictive indeed, at least for the narrowest version of the language, the one suitable, as he says, for 'limning the true and ultimate structure of reality' (Quine 1960, 221), or setting down 'all traits of reality worthy of the name' (Quine 1960, 228). That language would, for example, exclude the general idiom of indirect discourse, subjunctive conditionals, and indexical terms (see, for example, Quine 1960, sections 45–6). A string of symbols which irreducibly contains an expression which does not occur in the imagined scientific language is, of course, not a sentence of that language: in that language, it is meaningless.

In some cases, Quine appeals to this kind of notion of meaninglessness. The clearest example, perhaps, occurs in a discussion of the underdetermination of theory by evidence. Quine suggests that we might respond to the existence of a global theory empirically equivalent to our own not by accepting that theory and our own as both true but rather by counting the 'irreducibly alien terms' of the other theory 'as meaningless' (Quine 1990a, 98) and excluding them from the vocabulary of our scientific language. The sentences of the scientific language which use such terms then count as meaningless simply because they include terms which are not in the vocabulary of that language; no philosophically interesting or controversial notion of meaninglessness is involved.

Both in cases like indirect discourse and in the case of the imagined rival theory, what's doing the philosophical work is not the notion of meaninglessness but rather the decision to exclude certain terms from the scientific language. In the former kind of case, certain idioms of ordinary language are excluded from scientific language. That language is designed for maximum clarity and objectivity; idioms are excluded because they fail to meet these requirements, not because they are meaningless as idioms of ordinary language. In the 1977 essay, 'Facts of the Matter', Quine says:

> Ordinary language is only loosely factual, and needs to be variously regimented when our purpose is scientific understanding … We withdraw to a language which … is visibly directed to factual distinctions—distinctions that are unquestionably underlain by differences, however inscrutable, in elementary physical states. (Quine 1977, 168)

Failing to meet the standards for scientific language does not imply or even suggest that a given idiom of ordinary language is meaningless. To the contrary: Quine explicitly recognizes that excluded idioms may be indispensable for legitimate purposes other than 'limning the true and ultimate structure of reality'. In 'Facts of the Matter', he says: 'I do not advise giving up on ordinary language, not even mentalistic language' (Quine 1977, 168). In a discussion of the reasons for excluding indirect discourse from the most austere version of canonical notation, Quine says that the idiom is nevertheless not 'humanly dispensable' (Quine 1960, 218). After emphasizing how much of our ordinary language is excluded by canonical notation he says: '[n]ot that the idioms thus renounced are supposed to be unneeded in the market place or the laboratory' (Quine 1960, 228), and goes on to say excluded idioms may be needed to teach some of those which are accepted as part of canonical notation. In 'The Scope and Language of Science' he says explicitly that scientific language is 'a splinter of ordinary language, not a substitute' (Quine 1957, 236). The limits of Quine's scientific language are clearly not supposed to correspond to the limits of meaningful ordinary language; on Quine's view there are two wholly different issues here.

The scientific language will thus exclude some sentences simply because they contain terms which are not present in that language. Within the language, however, Quine makes no effort to separate grammatical sentences into the meaningful and the meaningless; indeed he sees no reason to try. In *From Stimulus to Science*, he defines the empirical content of a set of sentences as, roughly, the observation categoricals that it implies. (An observation categorical is a standing sentence—true or false once for all, rather than having a truth-value that changes from occasion to occasion—which is directly refutable by observation sentences; see e.g. Quine 1995, 25, 43ff.) He speaks of a set of sentences which has empirical content as having 'critical mass' (Quine 1995, 48). Many sentences which play an undeniable role in our empirical knowledge have, by this definition, no empirical content at all if we consider them individually, in isolation from other sentences with which they interlock. (This is one way of stating holism.)

Quine considers the natural suggestion that we can attribute empirical content to an individual sentence if removing it from a set of sentences results in a new set which does not imply all of the observation categoricals implied by the old set. He rebuts this suggestion with a well-known argument which shows that this criterion would be useless, excluding nothing.[7] Consider an arbitrary sentence (or sequence of words, at least); call it *q*. Take some arbitrary observation categorical, *c*. Now consider the set containing *q* and the sentence 'If *q*, then *c*.'. This set has empirical content, since it implies the observation categorical, *c*; removing *q* from the set results in a set that has no empirical content. Hence, by the proposed criterion, *q* would count as having empirical content; but *q* was an arbitrary sequence of words. Quine comments on this argument as follows:

> The sentences we would like to credit with empirical content are ones that are supporting members of *interesting* sets with critical mass, sets that are not only testable but worth testing … I see no way of molding this requirement into a rigorous standard of shared content. The really clear notion of having content is just critical mass. Some single sentences have it, most do not. (Quine 1995, 48)

The implication here is that having critical mass is the *only* really clear notion of having content. Given that most sentences—in particular, most sentences of serious science—do *not* have critical mass, lacking critical mass is clearly not the same as being meaningless. So, Quine implies, there is no clear notion of meaninglessness.

Strikingly, Quine makes it clear that even if there were a workable criterion of meaninglessness to be had along these lines, he would not want to apply it within canonical notation:

> Even if I had a satisfactory notion of shared content, I would not want to impose it in a positivist spirit as a condition of meaningfulness. Much that is accepted as true or plausible in the hard sciences, I expect, is accepted without thought of its joining forces with other plausible hypotheses to form a testable set … Surely it often happens that a hypothesis remote from

7. A version of this argument occurs in the Preface to the second edition of A. J. Ayer, 1936/1946, 11f. Cf. also Hempel, 106f.

> all checkpoints suggests further hypotheses that are testable … Positivistic insistence on empirical content could, if heeded, impede the progress of science. (Quine 1995, 48f.)

The same point occurs when Quine considers the idea (which he attributes to Natuhiko Yoshida) that 'our broadest scientific laws' may 'escape evidence altogether' (Quine 1990b, 13). He comments: 'Very well: if in a final full accounting this is the state of affairs, then those laws must be accounted devoid of empirical content on any reckoning, mine or another' (*loc. cit.*). Again, it is clear that he does not want to exclude sentences which lack empirical content: even if a verificationist criterion of meaningfulness were available, he would not want to impose it on our theory. Quine also explicitly accepts that mathematical sentences, and even the set of all such sentences, likewise lack empirical content (see Quine 1995, 53.) But clearly he does not think of them as meaningless, and neither does he follow the Logical Empiricists in thinking of them as analytic.

In similar vein, Quine rejects the idea of imposing categorial distinctions to limit what counts as a meaningful sentence of canonical notation. He acknowledges that 'there has been a concern among philosophers to declare meaningless, rather than trivially false, such predications as "This stone is thinking about Vienna" (Carnap) and "Quadruplicity drinks procrastination" (Russell)' (Quine 1960, 229). He comments:

> But since the philosophers who would build such categorial fences are not generally resolved to banish from language all falsehoods of mathematics and like absurdities, I fail to see much benefit in the partial exclusions that they do undertake; for the forms concerned would remain still quite under control if admitted rather, like self-contradictions, as false.... Tolerance of the don't-cares … is a major source of simplicity of theory … (*loc. cit.*)

These are clearly not the words of a philosopher who wants to use the idea of meaninglessness to do serious philosophical work.

VI.

To sum up very briefly: the idea of meaninglessness—nonsense, if you like—plays a significant role in the history of analytic philosophy. My concern here has been with the fate of that idea at the hands of Quine. My conclusion is that there is simply no idea of nonsense that is defensible in Quinean terms which will play any significant philosophical role. Given Quine's overarching naturalism, there is no position from which the philosopher can tell otherwise competent speakers of the language that a given sentence of theirs is nonsensical. This assessment might seem to be called in question by the favourable view of the notion of meaningfulness in 'The Problem of Meaning in Linguistics'. I have argued, however, that consideration of the basis of that favourable view shows that the essay is in fact an endorsement of the idea: whether a sentence is meaningful is simply a matter of whether it is in use, or useable, within the community of those who speak the given language, so no sentence which speakers of the language use can be dismissed as meaningless.

Quine's work contains a number of favourable references to verificationism, and this has led some to attribute to him a verificationist criterion of meaningfulness, albeit a version of that criterion modified by holism. He may indeed have been attracted, for a time, by a version of verificationism of this sort, and may even have thought that such a view followed from his naturalistic account of language acquisition. (Whether Quine ever really believed that is, I think, unclear.) Here I have argued that no version of the verificationist criterion can be justified on the basis of Quine's general views, and that those views do not support the idea of any clear-cut distinction between the meaningful and the meaningless (except perhaps the altogether anodyne behavioural idea mentioned in the previous paragraph).

In the context of Quine's idea or ideal of a scientific language, finally, I have also argued that the idea of meaninglessness has no real application. Some terms of ordinary language will be excluded from the scientific language simply because they lack sufficient clarity and objectivity. Sentences using those terms are, of course, not meaningful sentences of the language from which the terms are excluded, but this no more involves a philosophically interesting or controversial sense of 'meaningless' than does the fact that a sequence of German words is not a meaningful sentence of French. Within the scientific language, Quine explicitly rejects the idea that we should attempt to demarcate which superficially grammatical sequences of its words form meaningful sentences and which do not; the issue is simply of no concern to him.[8]

8. I should like to thank the audiences at two conferences where I read versions of this paper: the second annual conference of the Society for the History of Analytic Philosophy and the Leonard Linsky Memorial Conference. For their comments on earlier versions, I am also indebted to Dirk Greimann and, especially, to Andrew Lugg.

References

Ayer, Alfred J. 1936/1946: *Language, Truth, and Logic*. London: Gollancz, 1936. Second edition 1946.

Carnap, Rudolf 1928/1967: *Scheinprobleme in der Philosophie*. Berlin: Weltkreis. Translated by Rolf A. George as *Pseudoproblems in Philosophy*. Berkeley, Los Angeles: University of California Press.

— 1934/1937: *Die Logische Syntax der Sprache*. Vienna: Julius Springer Verlag. Translated by Amethe Smeaton as *Logical Syntax of Language*. London: Kegan Paul Trench, Trubner & Co., 1937.

— 1950/1956: "Empiricism, Semantics, and Ontology." *Revue Internationale de Philosophie* 4, 20-40. Reprinted in: Rudolf Carnap, *Meaning and Necessity*. Chicago: University Press of Chicago, second edition, 1956.

— 1963: "Autobiography". In: Paul A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*. LaSalle, IL: Open Court, 3–84.

Dummett, Michael 1974/1978: "The Significance of Quine's Indeterminacy Thesis". *Synthese* 27, 351–97. Reprinted in: Michael Dummett, *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press, 375–419.

Gibson, Roger F. Jr. 1988: *Enlightened Empiricism*. Tampa, FL: University of South Florida Press.

Hempel, Carl G. 1965: "Empiricist Criteria of Cognitive Significance". In: Carl G. Hempel, *Aspects of Scientific Explanation*. New York: Free Press, 101–119.

Quine, Willard V.O. 1953a: "Two Dogmas of Empiricism". In: Willard V.O. Quine, *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 20–46.

— 1953b: "The Problem of Meaning in Linguistics". In: Willard V.O. Quine, *From a Logical Point of View*. Cambridge, MA: Harvard University, 47–64.

— 1957: "The Scope and Language of Science". *British Journal for Philosophy of Science* 8, 1–17. Reprinted in: Willard V.O. Quine, *Ways of Paradox*. Cambridge MA: Harvard University Press. Expanded edition 1976, 228–245.

— 1960: *Word and Object*. Cambridge, MA: MIT Press.

— 1963/1976: "Carnap and Logical Truth". In: Paul A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*. LaSalle, IL: Open Court, 385-406. Reprinted in: Willard V.O. Quine, *Ways of Paradox*. Cambridge, MA: Harvard University Press. Expanded edition 1976, 107–132.

— 1966: "Limits of Knowledge". In: Willard V.O. Quine, *Ways of Paradox*. Cambridge, MA: Harvard University Press, 59–67.

— 1969: "Epistemology Naturalized". In: Willard V.O. Quine, *Ontological Relativity and Other Essays*. New York: Columbia University Press, 69–90.

— 1970a: *Philosophy of Logic*. Englewood Cliffs, NJ: Prentice Hall.
— 1970b: "Philosophical Progress in Language Theory". *Metaphilosophy* 1, 2–19.
— 1974: *Roots of Reference*. La Salle, IL: Open Court.
— 1975: "The Nature of Natural Knowledge". In: Samuel Guttenplan (ed.), *Mind and Language*. Oxford: Oxford University Press, 67–82.
— 1977: "Facts of the Matter". In: Robert Shahan (ed.), *American Philosophy from Edwards to Quine*. Norman, OK: University of Oklahoma Press, 176–196.
— 1986: "Reply to Roger F. Gibson Jr.". In: Edwin Hahn & Paul Arthur Schilpp (eds.), *The Philosophy of W. V. Quine*. Chicago and LaSalle, IL: Open Court. 1986. Expanded edition 1998, 155–157.
— 1990a: *Pursuit of Truth*. Cambridge, MA: Harvard University Press.
— 1990b: "Three Indeterminacies". In: Robert B. Barrett & Roger F. Gibson (eds.), *Perspectives on Quine.* Oxford: Blackwell, 1–16.
— 1991: "Two Dogmas in Retrospect". *Canadian Journal of Philosophy* 21, 265–74.
— 1995: *From Stimulus to Science*. Cambridge, MA: Harvard University Press.
Raatikainen, Panu 2003: "Is Quine a Verificationist?". *The Southern Journal of Philosophy* 41, 399–409.
Trillin, Calvin 1978/2007: *Alice, Let's Eat*. New York: Random House. Random House Trade Paperback Edition 2007.

Grazer Philosophische Studien
89 (2014), 135–152.

ARE THERE EMPIRICAL CASES OF INDETERMINACY OF TRANSLATION?

Rogério Passos SEVERO
Universidade Federal de Santa Maria

Summary

Quine's writings on indeterminacy of translation are mostly abstract and theoretical; his reasons for the thesis are not based on historical cases of translation but on general considerations about how language works. So it is no surprise that a common objection to the thesis asserts that it is not backed up by any positive empirical evidence. Ian Hacking (1981 and 2002) claims that whatever credibility the thesis does enjoy comes rather from alleged (fictitious) cases of radical *mis*translation. This paper responds to objections of that kind by exhibiting actual cases of indeterminacy of translation.

Introduction

"Gavagai" is a made up word, as are the various translations Quine says it admits—all equally compatible with the behaviors of the made up native speakers, but incompatible with each other. Apparently some of the most impressive cases of indeterminacy of translation are fictitious. My favorite is Jorge Luis Borges's (1964) description of a few pages of the "Eleventh Volume of *A First Encyclopedia of Tlön.*" We are told that the languages spoken on planet Tlön differ radically from the ones we speak on Earth. Commenting on one of those languages, Borges produces a vivid image of what indeterminacy of translation might look like. He mentions a native sentence and two possible translations. One of them is a literal translation; the other has a more natural expression in most human languages:

> The nations of this planet are congenially idealist. […] The world for them is not a concourse of objects in space, but a heterogeneous series of independent acts. It is successive and temporal, not spatial. There are no nouns in Tlön's conjectural *Ursprache*, from which the "present" languages and dialects are derived […]. For example: there is no word corresponding to the

> word "moon," but there is a verb which in English would be "to moon" or "to moonate." "The moon rose above the river" is *hlor u fang axaxaxas mlo*, or literally: "upward behind the onstreaming it mooned." (Borges 1964, 8)

This gives us an illustration of indeterminacy of translation because the native sentence *hlor u fang axaxaxas mlo* apparently can be equally well translated as "The moon rose above the river" and as "Upward behind the onstreaming it mooned." The latter is a more literal translation, but harder for us to understand. The former is easier for us to understand, but might render other portions of the native discourse less readily intelligible: portions of their philosophical and scientific discourse might sound nonsensical to us when translated out of their native idiom. Thus, alternative manuals of translation—one more literal, the other less so—might be thought up which afford *roughly equal fluency* in dialogues and negotiations with the natives of Tlön, but which diverge in the translation of individual sentences. I take this to be a good illustration of Quine's thesis of the indeterminacy of translation.[1]

The question addressed in this paper is whether there is any actual empirical evidence for the thesis. A recurrent objection says that there is none, and that the thesis asserts a mere logical possibility.[2] Quine himself did not do much to prevent this kind of objection from coming up. His reasoning contains very little in terms of direct positive evidence.[3] It relies instead on considerations about how language and translation works in general, not on actual case studies. Quine does not think that the lack of direct evidence counts against the thesis. Rather, he argues that this is to be expected, given how hard it usually is to find a single manual of translation.[4] Once a manual of translation is found that affords fluency in

1. This is how Quine formulates the thesis in *Word and Object*: "manuals for translating one language into another can be set up in divergent ways, all compatible with speech dispositions, yet incompatible with one another" (1960, 27). The notion of 'incompatibility' that figures in this passage is explained in *Pursuit of Truth* in terms of non-interchangeability: the "two translation relations might not be usable in alternation, from sentence to sentence, without issuing in incoherent sequences. Or, to put it in another way, the English sentences prescribed as a translation of a given […] sentence by two rival manuals might not be interchangeable in English contexts" (Quine 1992a, 48).

2. Collin and Guldman (2005, 255), for example, say that "… it remains a striking feature of his account that Quine only argues for the abstract logical possibility of the indeterminacy of translation. He never offers serious examples taken from actual anthropological or linguistic research." See also Bar-On (1993), who argues that indeterminacy of translation is inconsistent with our actual translation practices, and Hacking (1981 and 2002), discussed below.

3. See Quine (1960, chapter 2), (1970), and (1987).

4. "Radical translation is a rare achievement, and it is not going to be undertaken success-

dialogues and negotiations, why keep on seeking for another? Although Quine's reasoning does not require direct empirical evidence, I shall argue his thesis is confirmed by some case studies in radical translation. The cases presented below are offered here also as a response to the claim that indeterminacy of translation is a mere logical possibility that has little to do with our actual translation practices. This suggestion has appeared several times in the literature, but perhaps its most striking appearance has been in Hacking's (1981).

Hacking argues that because Quine's reasons for indeterminacy are so abstract and theoretical, whatever empirical credibility the thesis has must come instead from a few notorious cases of radical *mistranslation*. He then shows that these cases are all fictitious, and concludes it is unlikely that there has ever been a case of radical mistranslation. Hacking seems to suggest that cases of mistranslation offer evidence for indeterminacy of translation. But surely this is not how Quine viewed the matter: indeterminacy of translation says that if a translation manual can be devised, so can others that are equally compatible with the behaviors of the natives but incompatible with each other. So it is a thesis about multiple translatability, not about untranslatability or mistranslatability. Hacking's point, however, is that given the lack of direct empirical support for indeterminacy of translation, it might gain some plausibility from cases of mistranslation; but—and this is his main argument in (1981 and 2002)—the allegedly historical cases of mistranslation are all fictitious. This does not entail that indeterminacy is impossible, but it is meant to drain most of its plausibility. Hacking concludes that given the lack of empirical support, indeterminacy of translation is a logical possibility (something we cannot prove impossible) that is most likely false of the world we live in.[5] Given Hacking's argumentative strategy, the bulk of his reasoning turns on an analysis of three notorious cases of alleged radical mistranslations. One of these cases is that of an alleged mistranslation of the word 'kangaroo':

> On their voyage of discovery to Australia a group of Captain Cook's sailors captured a young kangaroo and brought the strange creature back on board

fully twice for the same language" (Quine 1992a, 50f.).

5. In his *Historical Ontology* (2002, 152) Hacking added a few extra sentences at the very beginning of his (1981) paper, which is reprinted in the book: "Some readers will protest that this shows nothing about Quine's logical point. I am not so sure. If something is claimed as a logical possibility about translation, which is never known to be approximated for more than a few moments in real life, may we not begin to suspect that the conception of translation that is taken for granted may be erroneous?"

> their ship. No one knew what it was, so some men were sent ashore to ask the natives. When the sailors returned they told their mates, 'It's a kangaroo.' Many years later it was discovered that when the aborigines said 'kangaroo' they were not in fact naming the animal, but replying to their questioners, 'What did you say?'[6]

As Hacking points out, this report is false. In the Guugu Yamidhirr dialect, spoken by Aborigines who lived in the area where Cook landed, the word for kangaroo is "ga*n*urru", where "*n*" is a phoneme that sounds a bit like "ng." According to Hacking (1981, 172), this was "apparently pointed out in a letter to an Australian newspaper in 1898," but only became common knowledge with the work of anthropologist John Haviland in 1972. Travelers in Australia subsequent to Cook apparently either failed to contact speakers of the Guugu Yamidhirr dialect or made contact but failed to pronounce the word properly; hence the myth of the radical mistranslation of "kangaroo." There was no mistranslation, just poor phonetic transcription. Two other cases of alleged mistranslation are likewise analyzed away by Hacking—that of the French word "vasisdas" and that of the English word "indri". Based on his analysis of these cases, Hacking suggests that there is no evidence of there ever having occurred a single case of radical mistranslation.

This paper does not examine the examples brought by Hacking—which are indeed fictitious—but discusses instead cases of radical translations of Amerindian words and phrases that apparently satisfy Hacking's definition of a mistranslation (section 1). Amerindian cosmologies—found especially in native cultures of the Amazon region, but also throughout North, Central and South America—are so much at odds with the cosmologies prevalent in Europe (and throughout the world nowadays) that radical mistranslations in Hacking's sense are bound to occur. I then argue (section 2) that there is something wrong with Hacking's criteria of mistranslation, and that the cases exhibited here are in fact evidence both of indeterminacy of translation and of what one might want to call 'cosmological relativity'. The paper concludes that indeterminacy can be argued for using both top-down (from abstract reasons, as Quine did) and bottom-up (from actual cases of translation, as we do here) strategies, and at the very end answers a couple of objections.

6. Quoted from Hacking (1981, 175), originally in *The Observer* (London, 1973). See also Banks's entry 14, July 1770, in his (1962).

1. *Radical "mistranslations" of Amerindian phrases*

We begin with Hacking's criteria for a radical mistranslation:

> (1) Speakers of two very different languages are trying to communicate. (2) A speaker of one language says *s*. Speakers of the other language take him to be saying *p*. (3) This translation is completely wrong. Yet (4) neither party realizes it, although they continue to converse. Moreover (5) the mistranslation persists until it is too late to correct. (1981, 171)

Hacking has in mind cases of mistranslation of names, hence cases of *malostension*, or the misidentification of the object or objects referred to by a name. These occur "when (6) an expression of the first language is taken by speakers of the second language to name a natural kind. (7) It does nothing of the sort, but (8) the second language incorporates this expression as the name of the natural kind in question" (171). Conditions (3) and (7) are meant to rule out "mere differences in nuance, moderate misunderstandings and misclassifications [...], or the taking of the name of an individual as the name of a class" (171). As we shall see next, conditions (1)–(8) are apparently satisfied by some translations of Amerindian words.

Anthropological studies have been pointing out for some time now that most Amerindian peoples do not conceive themselves as the only creatures that see themselves as humans.[7] Like many other cultures, they describe themselves as persons and as human beings, and they also conceive persons as centers of intentionality and agency. But, unlike many cultures, they view the belonging of an individual to a natural kind as something quite different from what we take it to be. For many cultures, this is a matter of having certain natural traits (biological, physical, etc.) which are true of the individuals of that kind regardless of how they are perceived by others. For the Amerindians, on the other hand, belonging to a natural kind is a matter of perspective. The same individual that from a human perspective is a jaguar, is said to be a human being from the perspective of the jaguars (see Lima 2005, 215), and is said to belong to yet another kind from the perspective of other creatures (say, a fish, an armadillo, a monkey, a spirit, or whatever). In other words, for the Amerindians the natural sorting of an individual turns on the species that sees that individual. The kind to which an individual belongs is relative to how it is seen by others. So Amer-

7. See, e.g., Århem (1993), Descola (1996), Lima (1996, 2005), Viveiros de Castro (1996, 1998, 2002), and Vilaça (2005).

indians, like other cultures—European cultures, for example—claim that human beings are persons. But they diverge from others in saying that not only we, humans, see ourselves as humans. They say that seeing oneself as human is a common trait of all creatures, whereas others—Europeans, for example—would tend to say that humanity is what sets us apart from other creatures. They say that humanity is shared, and that what sets creatures apart is instead the kind of body that each has. On their view, the same individual that is a human from one perspective can also be non-human in another. Vilaça (2005, 450) describes the case of the Amazonian Wari' people and provides further references:

> Although they see jaguars as animals, the Wari' know from their shamans that jaguars see themselves as humans: that is, as people pursuing a full social life and endowed with a human appearance. A similar instance among the Carib of British Guiana, taken from Ahlbrinck's work of 1924, is cited by Levy-Bruhl as an example of this extended notion of humanity: "[A]nimals (just as plants and inanimate objects) live and act like humans. In the morning, the animals go 'to work,' as the Indians do. The tiger, the snake and all the other animals leave to go hunting; like the Indians, they must 'look after their family' …" (Ahlbrinck 1924, 221, in Levy-Bruhl 1996 [1927], 30).

Commenting on studies such as these, Viveiros de Castro (1998) offers a more generalized account of Amerindian cosmology:

> Typically, in normal conditions, humans see humans as humans, animals as animals, and spirits (if they see them) as spirits; however, animals (predators) and spirits see humans as animals (as prey) to the same extent that animals (as prey) see humans as spirits or as animals (predators). By the same token, animals and spirits see themselves as humans: they perceive themselves as (or become) anthropomorphic beings when they are in their own houses or villages and they experience their own habits and characteristics in the form of culture—they see their food as human food (jaguars see blood as manioc beer, vultures see the maggots in rotting meat as grilled fish, etc.), they see their bodily attributes (fur, feathers, claws, beaks, etc.) as body decorations or cultural instruments, they see their social system as organized in the same way as human institutions are (with chiefs, shamans, ceremonies, exogamous moieties, etc.). This 'to see as' refers literally to percepts and not analogically to concepts … (470)

Viveiros de Castro and others thus say that Amerindian cosmology has a "perspectival quality." Differences among kinds of creatures are not accounted for in physical or biological terms as many cultures understand them, but in terms of the perspective afforded by the body of the individual that perceives the individuals at hand. From the perspective of one's own body, one sees oneself as human, and sees other creatures as having different kinds of bodies, some of them non-human. But this is also true of the way all other creatures see themselves and the creatures around them.

One significant consequence of this is that in Amerindian tongues the words used to designate what we call "persons" or "humans", and which have been translated accordingly—e.g., *dene* (McDonnell 1984), *masa* (Århem 1993), *matsigenka* (Rosengren 2006), *wari'* (Vilaça 2005)—do not designate persons or humans as we understand them. Instead, those words function as pronouns or indexicals of self-designation—much like "we" or "us"—which vary in content according to who uses them and in which context. It is of course understandable that the Amerindian words just mentioned have been translated for "human beings" or "persons," and it is for us natural to continue to do so, given the fluency in dialogues and negotiations allowed by that choice. The same is true not only of words of self-designation but also of words that we usually translate as names of natural kinds such as jaguar, tapir, arapaima, etc. The Tupinambás (of eastern Brazil), for example, use the word *jauára*—also transcribed as *ya'guara*—to designate creatures of a natural kind (the jaguar), as we do, i.e. creatures that have a certain type of body. But for them, having that body is not something that belongs to a creature's independent nature or essence; rather it is something that a creature has or does not have relative to the perspective from which its body is perceived. The same individual creature may have the body of a jaguar when seen from the perspective of a human body, a body of a human being when seen from the perspective of a jaguar, and yet a different type of body from the perspective of a third creature. In fact, *jauára* works much like the Amerindian words for "person" and "human being": it registers a certain perspective, and functions much like an indexical, such as "you" or "they".

Reporting on his voyages to Brazil in the 16th Century, Hans Staden recalls being made captive by the native Tupinambás. He describes a ritual in which a Tupinambá declared himself to be a jaguar while eating human flesh: *jauára iche* ["I am a jaguar"] (see Staden 2008, 91). Being a *jauára* is in this case the perspective of a creature that eats human flesh, among other things, i.e. the perspective of a predator—but note that from that perspective it is not human flesh that is being eaten. The perspective one has is fixed by one's body, but bodies are in this framework essentially unstable and can change radically in special circumstances. Vilaça (2005) reports the case of a Wari' child who was invited by her mother to take a trip into the forest:

> Many days go by as they walk around and pick fruit. The child is treated normally by her mother until one day, realizing just how long they have spent away from home, the child starts to grow suspicious. Looking carefully, she sees a tail discreetly hidden between her mother's legs. Struck by fear, she cries for help, summoning her true kin and causing the jaguar to flee. (451)

Reports such as these are quite common in the Amazon region and offer evidence of how radically different from ours the notion of a body is for the Amerindians: it is not a substance or a physical substrate, but primarily a set of "affections or ways of being" (Viveiros de Castro 1996, 128), "a way of being actualized in a bodily form" (Vilaça 2005, 450).

What the Tupinambás and other Amerindians ordinarily see when they look at a jaguar is a jaguar, but this is not how jaguars see themselves, and neither is it what the Tupinambás and other Amerindians see in some special circumstances. The case is likewise for other creatures and even spirits and celestial bodies such as the moon (see Fernandes 1970, 171). This is so because the individuals that we may conceive as jaguars are conceived by the Amerindians as seeing themselves as humans; hence, they too have a language and designate themselves with words that correspond to *wari'*, *dene*, *masa*, *matsigenka*, etc. From their perspective, they also see beings that differ from themselves in bodily appearance, and are accordingly classified as predator or prey—just like we do with other creatures. Snakes and jaguars see themselves as humans, and in turn see humans as tapirs or white lipped peccaries, for example, as prey (see Baer 1994, 224, quoted by Viveiros de Castro 1998, 477). Hence, in Amerindian tongues the words we ordinarily translate as names of natural kinds, such as jaguar, tapir, armadillo, etc. vary in content (extension) according to who uses them and in which context, while not varying in what Kaplan (1989, 505ff.) calls "character." Like the Amerindian words for "human" or "person", they function as indexicals or pointers. They may of course be translated into many *other* languages as jaguar, tapir, armadillo, etc., and this is as good a translation as we will ever get without radical changes in our use of our words.

Strikingly, however, these translations satisfy Hacking's conditions for a radical *mistranslation:* a number of cultures have been in dialogue with Amerindian peoples over the last five centuries, translating words such as *jauára* for "jaguar". Yet what an English (or Portuguese, etc.) speaker means by "jaguar" differs radically from what the Tupinambás mean by *jauára*; and the difference here is not just a matter of nuance, moderate misunderstanding or misclassification, nor is it mistaking the name of an individual for the name of a class. In translation, "jaguar" ends up meaning a creature that belongs to a set picked out by their physical and biological traits independent of who sees it, whereas by *jauára* the Tupinambás mean a perspective which many individuals of different species can take on, including human beings. This difference, however, did not prevent the word "jaguar" and others like it from being incorporated from the Tupi language into Portuguese, Spanish, English, and other European tongues to designate a natural kind.[8] So we do have here historical cases of radical mistranslation in Hacking's sense.[9] At the same time, it is unclear which alternative translations would be better suited for these cases.

2. *Indeterminacy of translation and cosmological relativity*

The anthropologists mentioned above have pointed out that the cosmologies of the Amerindian peoples differ radically from ours, and they have offered indications on how Amerindians think (the inferences they make) and are inclined to talk on given occasions. For the most part they have not provided better translations, nor are they saying that the translations we do have are wrong. Instead, the suggestion is that in translation we are bound to use the categories with which we are familiar and project them onto native cultures. But this is precisely the point of Quine's indeterminacy of translation: "What the indeterminacy thesis is meant to bring

8. This happened with many other Tupi words as well. The Portuguese word for armadillo, for example, is "tatu", from the Tupi word *ta'tu*; "jaguatirica" (ocelot) comes from the Tupi *îaguara tyryk*; "guapuruvu" (s*chizolobium parahyba*) comes from ïwakuru'mbu.

9. Two interesting additional examples are those of the Wari' expressions *kwere-* and *jam-*, that are usually translated as body and soul (see Vilaça 2005, 452ff.). What we mean by our words "body" and "soul" has no counterpart in the Wari' cosmology. Having a *jam-* (soul) is for the Wari' having the capacity to transform, especially in extraordinary action. *Jam-* is not what gives a person's body feelings, thoughts, consciousness, etc., but what gives it its *instability.* A body—which is conceived by them not as a substance or substrate, but as a set of affections or ways of beings—will change due to its *jam-*. Vilaça reports (453) that "the Wari' insist that healthy and active people do not have a soul (*jam-*)", precisely because they are much less prone to change their affections or ways of being.

out is that the radical translator is bound to impose about as much as he discovers" (Quine 1992a, 49).

Hacking has argued previously for the *determinacy* of translation (see his 1975, 150ff.): he says that two translated sentences *p* and *q* cannot be both correct translations of a native sentence *s*, and at the same time be "contraries".[10] This is true, but it is a misreading of Quine's thesis, which does not say that the translated sentences *p* and *q* are contraries, but merely that the *manuals* that yield each sentence are incompatible in the sense that using them in alternation will bring about an incoherent English text or discourse: the translated sentences *p* and *q* are "not interchangeable in English contexts" (Quine 1992a, 48). The fact that they are not interchangeable does not entail that they cannot both be true. In fact, there is no reason to think that they are not *intertranslatable*: *p* is a translation of *s* which is a translation of *q*—thus *p* is a second-hand translation of *q*; moreover, they must both fit equally well the speech-behaviors of the natives while uttering *s*. So in many cases—though not necessarily in all—they are likely to have the same truth-value, and (by definition) are not contraries. The point is that even when two English sentences are perfectly good translations of *s*—in the sense of allowing for fluency in dialogues and negotiations—if we think of them as conveying meanings (conceived as something distinct from their actual behavior during those dialogues and negotiations), then the meanings of the two English sentences offered as translations of *s* must differ. If they did not differ, then they could for the most part be used in alternation without producing incoherence in the overall English text or discourse. So the fact that a translation relation is transitive—i.e. that if *p* is a translation of *s*, and *s* is a translation of *q*, then *q* is a translation of *p*—does not entail the transitivity of meaning. The sentences *p* and *q* might not be usable in alternation in English contexts, and thus there is hardly any sense in which they can be said to mean the same. In other words, a good translation is not evidence of sameness of meaning. Let's not dwell on Hacking's misreading of Quine's thesis here but merely press that the cases presented above are evidence of the *indeterminacy* of translation.

There are at least two ways of finding out what a native speaker of a foreign tongue means by what she says: we can translate her words into

10. It is not clear what Hacking means by saying that two sentences cannot be "contraries". Perhaps he means that they cannot be logically incompatible, i.e., that they cannot be negations of each other nor contraries in the strict sense (in which one says 'all S is P' and the other says 'no S is P')—in either sense the sentences cannot both be true.

a language we already know, or we learn to speak like her. In the latter case, little or no translation is needed. But in the former, finding out what a person means is the outcome of a translation; hence the meanings assigned to her words cannot be used as a standard for the correction of the translation itself. What we can do is to come up with a better translation—one that allows for more fluency in dialogues and negotiations—and with which the original translation can then be compared and corrected. In any case, if we are to say in our language what she says in hers, some translation will be needed.

To be sure, fluency in dialogues across cultures is bound to be broken here and there: some phrases will be untranslatable or only partially translatable. This surely happens with many Amerindian phrases in translation; and it is a common experience for anyone who speaks more than one tongue: one can know how to say things in a foreign language without ever quite finding a way of conveying it in one's mother tongue. The thesis of indeterminacy of translation has nothing to say about these cases. It is not a thesis about untranslatability, nor is it a thesis about mistranslations.[11] What it does say is that whenever we have a manual of translation that allows for dialogue and negotiations, *however broken*, then other manuals are possible that allow for *roughly equal* fluency in dialogues and negotiations, yet are incompatible with the original manual (in the sense mentioned above, that they cannot be used in alternation—that switching from one manual to another in the course of a translation will yield inconsistency in the translated text).

For the word *jauára* mentioned above, the standard translation is just "jaguar." This is of course the easiest and most natural way *for us* to understand what a Tupinambá says while pointing to a jaguar and uttering the word. But an alternative manual could try to be more faithful to what we now know about Amerindian cosmology by attempting to avoid projecting

11. Hacking (2002, 169) says that indeterminacy of translation "pulls in one direction and the idea of incommensurability"—which is usually defined in terms of untranslatability—"in the other". But here again Hacking's reading of the thesis of indeterminacy is mistaken: it assumes that the thesis entails that there are always "too many translations between schemes" (170). Yet indeterminacy is compatible with untranslatability, i.e. with there being no translation at all for a given set of sentences. And it is also compatible with there being only a few. Indeterminacy is one thing, translatability is another: "This thesis of indeterminacy of translation is by no means a theory of untranslatability. There are good translations and bad, and the two conflicting manuals imagined are good. However, there are also plenty of cases of untranslatable sentences, and they are commonplace even within our own language. A sentence about neutrinos admits of no translation into the English of 1900" (Quine 1992b, 1).

onto them our own theories about what a jaguar is. It could, for example, translate words such as *jauára* for phrases containing "jaguar-perspective" or "jaguar-from-our-perspective" or something of the kind. This would increase the intelligibility *for the Amerindians* of what we say, but at the cost of making what they say less readily intelligible to us. The standard translation (where *jauára* is just "jaguar"), on the other hand, projects our view of what a jaguar is onto the natives, and thus makes it harder for them to understand what we say, but easier for us to speak to them. So in choosing one manual over another, there is a trade off. To be sure, some translations are just wrong, in that the manuals that issue them systematically yield sentences that are incompatible with the speech behaviors of the natives. But the possibility of more than one manual issuing sentences that allow for dialogues and negotiations that are roughly equal in fluency seems to be implied by the differences of our own cosmology and that of the Amerindians. The question of whether by *jauára* the natives really mean "jaguar" or "jaguar-from-our-perspective" is in fact a question about which manual of translation is to be favored. If the manuals that issue them do in fact allow for roughly equal fluency in conversations, and if no other manual is available that allows for increased fluency, then there is hardly any sense in saying that only one of them captures what the natives really say. If translation according to one manual is correct, then so is the alternative. This is not to say that the natives do not know what they mean: certainly they know what they mean just as much as we do. By "jauára" they mean *jauára*, just as we mean jaguar by "jaguar". Surely there are occasions in which people do not know what they mean, and we might even want to say that meaning in these cases is indeterminate. But this is not the thesis of indeterminacy of translation.

Lima (1996, 30) describes the initial strangeness to her ears of certain Tupi phrases (spoken by the Jurunas, of the Amazonian lowlands) like *amãna ube wï*, literally: "it rained for me". She reports that most of the statements made by the Jurunas have the qualification "for me": it is beautiful *for me*, it turned into a jaguar *for me*, it is true *for me*, etc. For the Jurunas, however, it would make little sense to speak as we do, as if from nowhere. For the purposes of translation, of course, we could just say that *amãna ube wï*, said by a Juruna, is what we mean by "it rained," or "it rained where I was". But in doing so we eventually have to add in some explanation about why they seem to believe in claims that to us are obviously false or senseless, such as "this is blood for me but manioc beer for a jaguar", "while hunting he appeared as a pig to his friends, who then killed him", etc. Alternatively, we might try a translation that already has that "perspectival quality" built into it, thus allowing for a more literal rendering of sentences such as *amăna ube wï*: it rained for me. In this case it is the translated sentence itself that is harder for us to understand. So, again, in choosing one manual over another, there is a trade-off; and the fact that there is a trade-off is evidence of indeterminacy.

Although we have been speaking here of *the* thesis of indeterminacy of translation, it is in fact a set of theses containing at least two. This was not clear in Quine's earlier writings on the matter but gradually became more transparent. Quine came to speak of the indeterminacy of translation of *sentences* (or holophrastic indeterminacy), as distinguished from the indeterminacy of translation of subsentential parts, especially the indeterminacy of translation of *terms* (or indeterminacy of reference). The latter thesis admits of a proof, with the use of proxy functions:

> A proxy function is any explicit one-to-one transformation, *f*, defined over the objects in our purported universe. By 'explicit' I mean that for any object *x*, specified in an acceptable notation, we can specify *fx*. Suppose now we shift our ontology by reinterpreting each of our predicates as true rather of the correlates *fx* of the objects *x* that it had been true of. Thus, where '*Px*' originally meant that *x* was a *P*, we reinterpret '*Px*' as meaning that *x* is a *f* of *P*. Correspondingly for two-place predicates and higher. Singular terms can be passed over in view of §10.[12] We leave all the sentences as they were, letter for letter, merely reinterpreting. The observation sentences remain associated with the same sensory stimulations as before, and the logical interconnections remain intact. Yet the objects of the theory have been supplanted as drastically as you like. (Quine 1992a, 31f.)

This reasoning for the indeterminacy of reference came to be favored by Quine over the "gavagai" argument used in *Word and Object*, because it can be fleshed out into a full logical proof. This is in stark contrast with the stronger thesis of the indeterminacy of translation of *sentences*, for which there is no proof (see Quine 1992a, §§ 13 and 20). In his later writings Quine comes to describe it as a conjecture.[13] However that may be, both theses have implications for metaphysics. Indeterminacy of reference has

12. In §10 Quine describes a method for eliminating singular terms in favor of definite descriptions. This is essentially Russell's technique, but now extended to all singular terms. This is not to be understood as an interpretation of singular terms—i.e., it does not say or clarify what they mean—nor is it meant to replace singular terms in ordinary or scientific discourse. (Quine 1992a, 25–28)

13. See Quine (1998, 728); for further comments and discussion, see Hylton (2007, chapter 8).

well-known implications for the status of ontology, explicitly drawn by Quine himself in "Ontological Relativity" (1969). Holophrastic indeterminacy has an implication that has been less explicitly explored, which we might want to call "cosmological relativity". Whereas *ontological relativity* states that all existence claims are relative to a manual of translation, *cosmological relativity* says that all claims about the relations among entities are relative to a manual of translation.

The fact that Amerindian cosmologies have the "perspectival quality" described above, whereas other cosmologies do not, suggests cosmological relativity. In our cosmologies the attributes assigned to an individual do not turn on who is describing that individual. Hence, being objective usually means to describe or explain something without letting the particular perspective from which the description is made intrude. In Amerindian cosmologies, by contrast, the attributes assigned to an individual vary according to the bodily perspective from which it is perceived. Hence, the ideal of an objective view from nowhere is out of question. Objectivity is granted, rather, by seeing things from the perspective of the individual that is being described. To know a jaguar objectively is to become acquainted with its perspective, to see the world as it sees it, and so on.[14] These differences are so radical and run so deep that translations from Amerindian into European languages are bound to be quite loose at some points. The radical translator may opt for projecting more or less of his own cosmology into what is said. And this, we conclude, is evidence suggestive of cosmological relativity. Even in cases where the individuals of which Europeans and Amerindians speak can be matched up onto one another, they are conceived in radically different ways. Hence *cosmological* relativity can obtain even if *ontological* relativity does not. In translation, both might be suggested, but the Amerindian cases mentioned above are evidence primarily of the former.

3. *Two objections*

(1) Given the anthropological evidence presented above, one might want to say that we do in fact have good reasons for translating Amerindian

14. This in part explains why Amerindians were apparently so easily converted into Western religions, and also why they would so easily fall back into their own rituals. "Professing" the new faith was their way of finding out what it was about; the Europeans, however, mistook this as evidence of faint-heartedness (see Viveiros de Castro 2002, chapter 3).

sentences more literally: we have good anthropological evidence of how they think and what they believe in, and we should translate them accordingly. "*Amãna ube wi*" would then really mean "It rained for me" and not "It rained where I was."—This is an interesting objection, because it is indeed true that we have good anthropological evidence of a perspectival cosmology among Amerindians. This in turn provides clues as to which translations are empirically adequate. But it does not rule out alternative translations which have *roughly* equal adequacy. There is an issue here as to what exactly is to count as part of a translation: is it just the sentences translated, or those sentences and whatever else a translator might add so as to facilitate its understanding—say, footnotes, introductory remarks, explanations, gestures, etc.? "It rained for me" can only work as a good translation of "*Amãna ube wi*" when offered in a context of an explanation of how the Amerindians behave and think (why they are inclined to speak the way they do). So several other sentences have to be added to the translation so as to make it intelligible to us and usable in conversations with the natives. These other sentences include remarks to the effect that the Jurunas say "*Amãna ube wi*" in contexts where we would most likely just say "It rained where I was." These other remarks link the literal translation of the original sentence with sentences that are idiomatic in our own tongue. In effect we have here layers of translation. The translation "It rained for me" is at an intermediate level, between "*Amãna ube wi*" and "It rained where I was." The choice is then not between "It rained for me" and "It rained where I was", but between "It rained for me" plus an explanation of how we can understand this in our terms and "It rained where I was" plus an explanation of why this is not what the Jurunas literally say. The translated sentence itself is not the same in each case, but both alternatives will afford roughly equal fluency in dialogues in negotiations, given the explanatory remarks that accompany each translation. There is a trade-off between how much of the natives' views to build into the translated sentence and how much to convey by way of explanations and side remarks. So a translation which builds more of the natives' cosmological views into the translated sentences themselves and adds further remarks as to how to understand those sentences can do an equally good job of affording fluency in dialogues and negotiations as a translation that conveys less of the natives' cosmological views into the translated sentences but explains more of it in introductory and side remarks. Regarding which of these alternatives better captures what the natives really mean, there is indeterminacy: it is not something settled by our anthropological knowledge of the Amerindian cosmology; in fact, it is empirically irrelevant for anthropology. Both translation manuals seem equally compatible with our current anthropology.

(2) A related objection (see Hacking 2002, chapters 11 and 12) says the translation is not the issue in cases such as these. Rather, the difficulty lies in understanding the *style of reasoning* of the native speakers. The inferences they make are unlike the ones we make, as well as their ontology and cosmology.—This is in fact true, but why should this not render indeterminacy of translation even more plausible? If the style of reasoning of the natives differs radically from ours, then more introductory remarks, explanations and footnotes will be crucial to the understanding of the translated sentences. It is less likely that a straightforward single solution will clearly present itself as the translation of any given sentence. At least in some cases, it is likely that several translations will be roughly equally adequate, each accompanied by a different set of explanations, introductory remarks, footnotes, etc.

Conclusion

This paper has argued that Quine's original writings on the thesis of indeterminacy of translation can be supported by empirical evidence from actual cases of translation. Although Quine's views on the matter do not require direct evidence of the thesis—indirect, holistic considerations suffice—the fact that we can marshal some empirical support for this thesis fits nicely with Quine's empiricism and naturalism. Furthermore, it responds more straightforwardly to authors such as Hacking who take the lack of direct empirical support as evidence of the implausibility of the thesis. The paper has also indicated—without developing the point, however—that the thesis of ontological relativity, which is a direct consequence of the indeterminacy of reference, can be complemented with a thesis of cosmological relativity, which is a direct consequence of holophrastic indeterminacy. This is an issue that deserves further attention and has not been adequately handled here.[15]

15. Section 1 of this paper benefited substantially from an exchange with César Schirmer dos Santos, to whom I am most grateful. Many thanks also to Andrew Blom, Peter Hylton, Dirk Greimann, Ana Nicolino, Gilson Olegario da Silva, Jonatan Daniel, Laura Nascimento, Marcelo Fischborn, Tamires Dal Magro, André Abath, Ernesto Perini dos Santos, Mauro Engelmann, Flavio Williges, Ronai Rocha, Rogério Saucedo Corrêa, and Eros Carvalho.

References

Århem, Kaj 1993: "Ecosofia makuna". In: François Correa (ed.), *La selva humanizada: ecologia alternativa en el trópico húmedo colombiano*. Bogotá: Instituto Colombiano de Antropología, 109–126.

Banks, Sir Joseph 1962: *The Endeavour Journal 1768–1771*. Sydney: Angus & Robertson.

Borges, Jorge Luis 1964: "Tlön, Uqbar, Orbis Tertius". In: Jorge Luis Borges, *Labyrinths: Selected Stories and Other Writings*. New York: New Directions, 17–18.

Descola, Philippe 1996: *In the Society of Nature: A Native Ecology in Amazonia.* Cambridge: Cambridge University Press.

Fernandes, Florestan 1970: *A função social da guerra na sociedade tupinambá*. São Paulo: Livraria Pioneira.

Collin, Finn & Guldman, Finn 2005: *Meaning, Use, and Truth: Introducing the Philosophy of Language*. Hampshire, UK: Ashgate.

Hacking, Ian 1975: *Why Does Language Matter to Philosophy?.* Cambridge: Cambridge University Press.

— 1981: "Was there Ever a Radical Mistranslation?". *Analysis* 41(4), 171–175.

— 1983: *Representing and Intervening*. Cambridge: Cambridge University Press.

— 2002: *Historical Ontology*. Cambridge, MA: Harvard University Press.

Hylton, Peter 2007: *Quine*. New York: Routledge.

Kaplan, David 1989: "Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and other Indexicals". In: Joseph Almog, John Perry & Howard Wettstein (eds.), *Themes from Kaplan*. New York: Oxford University Press, 481–563.

Lima, Tânia 1996: "O dois e seu múltiplo: reflexões sobre o perspectivismo em uma cultura tupi". *Mana: Estudos de Antropologia Social* 2(2), 21–47.

— 2005: *Um peixe olhou para mim: o povo Yudjá e a perspectiva.* São Paulo: Editora UNESP.

McDonnell, Roger 1994: "Symbolic Orientations and Systematic Turmoil: Centering on the Kaska Symbol of *dene*". *Canadian Journal of Anthropology* 4, 39–56.

Quine, Willard V. O. 1960: *Word and Object*. Cambridge, MA: MIT Press.

— 1969: *Ontological Relativity and Other Essays*. New York: Columbia University Press.

— 1970: "On the Reasons for Indeterminacy of Translation". *Journal of Philosophy* 67(6), 178–183.

— 1987: "Indeterminacy of Translation Again". *Journal of Philosophy* 84(1), 5–10.

— 1992a: *Pursuit of Truth*. Rev. ed. Cambridge, MA: Harvard University Press.

— 1992b: “Commensurability and the Alien Mind”. *Common Knowledge* 1(3), 1–2.
— 1998: “Reply to John Woods”. In: Lewis E. Hahn & Paul. A. Schilpp (eds.), *The Philosophy of W. V. Quine*. 2nd exp. ed. Chicago: Open Court, 726–728.
Rosengren, Dan 2006: “Transdimensional Relations: On Human-spirit Relations in the Amazon”. *Journal of the Royal Anthropological Institute (N.S.)* 12, 803–816.
Staden, Hans 2008: *Hans Staden’s True History: An Account of Cannibal Captivity in Brazil*. Durham, N.C.: Duke University Press.
Vilaça, Aparecida 2005: “Chronically Unstable Bodies: Reflections on Amazonian Corporalities”. *Journal of the Royal Anthropological Institute (N. S.)* 11, 445–464.
Viveiros de Castro, Eduardo 1996: “Os pronomes cosmológicos e o perspectivismo ameríndio”. *Mana: Estudos de Antropologia Social* 2(2), 115–143.
— 1998: “Cosmological Deixis and Amerindian Perspectivism”. *Journal of the Royal Anthropological Institute (N. S.)* 4, 469–488.
— 2002: *A inconstância da alma selvagem*. São Paulo: Cosac Naify.

Grazer Philosophische Studien
89 (2014), 153–159.

SOME CRITICAL REMARKS ON QUINE'S THOUGHT EXPERIMENT OF RADICAL TRANSLATION

Oswaldo CHATEAUBRIAND
Pontifícia Universidade Católica do Rio de Janeiro/CNPq

Summary

Quine characterizes his argument for the indeterminacy of translation as a thought experiment, and claims that although it cannot be realized in practice, its result "is not to be doubted". Quine's thought experiment is a long argument, intended to have the character of a proof. The argument involves theoretical assumptions, such as the behaviorist approach—asserted by Quine to be "mandatory"—as well as empirical assumptions—e.g., that natives will have words for assent and dissent, which the linguist can recognize as such.

In this paper I critically examine some aspects of Quine's thought experiment and argue that even granting the behaviorist approach, the argument has substantial theoretical and empirical weaknesses for the claim of indeterminacy to be considered plausible.

1. *Introduction*

With the argument for indeterminacy of translation Quine intends to establish that there is no empirical basis for meaning and synonymy. In (Quine 1951) he argued strongly against the analytic-synthetic distinction, as based on meaning and synonymy, as well as against other notions which he classified as being part of what he called "the theory of meaning." At that time he drew a sharp distinction between the theory of meaning and the theory of reference, which included such notions as reference, extension, and truth, suggesting that whereas the latter had a sound empirical grounding, the former did not. The ensuing discussion was very lively, and many different issues were raised, but two responses—(Mates 1951) and (Carnap 1955)—seem to me particularly relevant to Quine's development of the indeterminacy argument.

Mates and Carnap argue that if what we consider as the empirical grounding of extension and reference is the observed use of words by speakers of a language, we can question the speakers with respect to the use of words in hypothetical and counterfactual situations to obtain inductively a sound empirical grounding for meaning (or intension).

Quine counters by elaborating the techniques suggested by Mates and Carnap into the notion of *stimulus meaning*, and argues that what this shows is not that meaning and intension have an empirical grounding, but that he, Quine, was wrong in assuming that reference and extension do. Herein lies the main conclusion of the indeterminacy of translation argument, which leads to a very substantial shift of perspective in Quine's later philosophy. A central aspect of the indeterminacy of translation depends on the argument that stimulus meaning is not sufficient to determine reference; hence the inscrutability of reference, which eventually leads to ontological relativity. The inscrutability of reference depends in turn on the indeterminacy of identity; i.e., in that stimulus meaning does not give us the relations of identity and diversity, which is nothing more than another way of saying that reference is socially inscrutable.

2. *Stimulus meaning*

Quine introduces the notion of stimulus meaning as follows (Quine 1960, 32f.):

> Let us make this concept of meaning more explicit and give it a neutrally technical name. We may begin by defining the *affirmative stimulus meaning* of a sentence … for a given speaker, as the class of all the stimulations (hence evolving ocular irradiation patterns between properly timed blindfoldings) that would prompt his assent. More explicitly, … a stimulation σ belongs to the affirmative stimulus meaning of a sentence S for a given speaker if and only if there is a stimulation σ' such that if the speaker were given σ', then were asked S, then were given σ, and then were asked S again, he would dissent the first time and assent the second. We may define the *negative* stimulus meaning similarly with 'assent' and 'dissent' interchanged, and then define the *stimulus meaning* as the ordered pair of the two … A stimulus meaning is the stimulus meaning of a sentence for a speaker at a date; for we must allow our speaker to change his ways. Also it varies with the modulus, or maximum duration recognized for stimulations. For, by increasing the modulus we supplement the stimulus meaning with some stimulations that were too long to count

before. Fully ticketed, therefore, a stimulus meaning is the stimulus meaning *modulo n* seconds of sentence S for speaker *a* at time *t*.

This notion of stimulus meaning, for an individual, appeals essentially to the totality of possible presentations of possible aspects of reality to that individual. The first 'possible' arises from the subjunctive form of Quine's definition. The second 'possible' arises from Quine's identification of the presentations with retinal images (in the visual case). This is a very strong notion of stimulation, which considerably sharpens the hypothetical and counterfactual questions suggested by Carnap and Mates. In fact, Quine is appealing to something akin to the notion of possible world (in the sense of counterfactual situations); and this is entirely natural because he is engaged in an impossibility argument. If we cannot get a workable notion of meaning from this strong apparatus, then we have a good justification for rejecting meanings—at least in the sense at issue according to which meaning has an empirical basis.

Stimulus meaning is generalized to the community of speakers to obtain an intersubjective notion of meaning, which yields a notion of intersubjective objectivity partly reflected in the notion of observation sentence. 'Red' as a one word sentence, meaning something like 'red there', has an objective meaning because the individual stimulus meanings agree to a very large extent over the whole community. It is an observation sentence partly because of that, and because this generalized agreement does not seem to depend on some additional generally shared information. 'Philosopher', also as a one-word sentence, meaning something like 'philosopher there', has a much less objective meaning, because there is likely to be a fairly large difference in the individual stimulus meanings depending primarily on collateral information.

3. *The thesis of the argument*

At the beginning of the chapter on translation and meaning, Quine formulates his thesis as follows (Quine 1960, 27):

> … the infinite totality of sentences of any given speaker's language can be so permuted, or mapped onto itself, that (*a*) the totality of the speaker's dispositions to verbal behavior remains invariant, and yet (*b*) the mapping is no mere correlation of sentences with *equivalent* sentences, in any plausible sense of equivalence however loose. Sentences without number can diverge

> drastically from their respective correlates, yet the divergences can systematically so offset one another that the overall pattern of associations of sentences with one another and with non-verbal stimulation is preserved. The firmer the direct links of a sentence with non-verbal stimulation, the less the sentence can diverge from its correlate under any such mapping.

He suggests, however, that it is better to formulate the thesis in terms of translation, and, more specifically, in terms of radical translation, where a linguist sets out to translate a language—say, Jungle—of which there is no previous knowledge and for which there are no interpreters. The purpose of the linguist is to develop a manual of translation from Jungle to English, say, which is essentially a function that to any sentence S of Jungle assigns a sentence S′ of English. Then, says Quine, the thesis is this (Quine1960, 27):

> … manuals for translating one language into another can be set up in divergent ways, all compatible with the totality of speech dispositions, yet incompatible with one another. In countless places they will diverge in giving, as their respective translations of a sentence of the one language, sentences of the other language which stand to each other in no plausible sort of equivalence however loose. The firmer the direct links of a sentence with non-verbal stimulation, of course, the less drastically its translations can diverge from one another from manual to manual.

The argument is then developed in great detail throughout the chapter and its main features are well known.

One problem with the argument, and not only with the argument but with Quine's work in general, is the assumption of behaviorism. Although this is already explicit in (Quine 1960), Quine makes an emphatic defense of it in (Quine 1987, 5):

> Critics have said that the thesis is a consequence of my behaviorism. Some have said that it is a *reductio ad absurdum* of my behaviorism. I disagree with this second point, but I agree with the first. I hold further that the behaviorist approach is mandatory. In psychology one may or may not be a behaviorist, but in linguistics one has no choice.

This is obviously a claim contested by many people, especially by Chomsky beginning in the 1950's (Chomsky 1959, 1969). It is now generally agreed that in linguistics it is not mandatory to be a behaviorist, nor can one distinguish behaviorism in linguistics and behaviorism in psychology. In any case, what Quine is acknowledging here is that the argument does depend on the assumption of behaviorism. It may not be a *reductio ad absurdum* of the assumption, but the assumption is present, and I take the argument as conditional on this assumption. It still does not take anything away from it's being a thought experiment about what can and cannot be done by behavioral methods.

Another problematic aspect of the methodology of the linguist is that in order to launch his inquiry he must be able to recognize assent and dissent—or agreement and disagreement—on the part of the native. Quine discusses this issue at some length (Quine 1960, 29f.), but his discussion is rather inconclusive, and after some general speculations, which do not settle the issue he remarks (1960, 30): "Let us then suppose the linguist has settled on what to treat as native signs of assent and dissent".

But how does the linguist know that the natives have signs of assent and dissent? This is supposed to be *radical* translation, after all, and the natives may have a very different attitude toward their environment than does the linguist. He points to a rabbit asking 'Gavagai?', and the native launches into a very long speech; he points to something else asking 'Gavagai?', and the native launches into another very long speech. No matter what he does, the result he gets is a very long speech. Perhaps there is something like agreement and disagreement in these speeches, but how is the linguist to determine that? Given Quine's methodology of stimulus meaning, it is absolutely essential that the natives behave in the way he envisions, but there is nothing the linguist can do if they don't. This is one reason why I think the approach through radical translation is problematic.

Yet another claim by Quine which seems to show the approach through radical translation to be problematic is made in (Quine 1987, 9), where he says:

> Radical translation is a near miracle, and it is not to be done twice to the same language. But surely, when we reflect on the limits of possible data for radical translation, the indeterminacy is not to be doubted.

Quine's first remark suggests a *practical* view of radical translation as having been done for the Jungle language and not to be done twice for that language. This is very misleading, however, because what we are considering is a thought experiment, and if one hypothetical linguist can produce a manual of translation from Jungle to English, then any number of hypothetical linguists can produce their own manuals as well. In fact, the point of switching from the first formulation of the thesis—in terms of a permutation of a speaker's language preserving dispositions to verbal behavior—to the formulation in terms of radical translation, is that if we have different manuals of translation, they can be used to generate such a permutation. Translate each sentence S of English to a sentence N of Jungle by the first manual, then re-translate N to a sentence S′ of English by the second manual and obtain the desired permutation by mapping S to S′. The problem is that we don't have the needed different manuals to do this, because, as a matter of fact, we have no idea what the Jungle language is like. It is just as hypothetical as the hypothetical linguists and the hypothetical natives, with no discernible structure.

The second remark suggests that the thought experiment has the quality of a proof, which we can see to be correct simply by reflection on the methods available to the linguist. Quine discusses some aspects of his methodology very thoroughly, but aside from some loose general remarks about pointing and the interpretation of reference, we have no clue as to the character of the possible different manuals for translating Jungle. In this sense I think the first formulation of the thesis in terms of a single speaker may yield a better approach than the formulation in terms of radical translation.

4. *Comparisons*

Gilbert Harman (1969, 14) compares Quine's thesis with the thesis that translation of number theory into set theory is indeterminate, because it can be done both using Zermelo's and von Neumann's methods—and Quine (1969, 296) basically concurs with him. This is a very superficial comparison, however, which gives no clue as to how it would apply to a natural language. And even if one were to do it for a more sophisticated theory, as in Gödel's arithmetization of metamathematics, which can be carried out in many different incompatible ways, it would still be only a superficial comparison with the claims in Quine's thesis.

A better approach may be to follow Richard Montague's later approach to the grammar of English. Montague (1973) wanted to defend the thesis that the structure of a natural language such as English is in a theoretical sense exactly the same as the structure of a formal language. In order to make this plausible he formulated a fragment of English for which he developed a grammar and an intentional semantics along the lines of a formal language. He did not treat the whole of the English language, of course, but gave an idea of how to proceed, and there was a substantial development of his ideas by a number of linguists. In order to make it plausible that permutations such as those suggested by Quine do in fact exist, one could similarly try to take a (non-trivial) fragment of English and define precisely how to map it onto itself satisfying conditions (*a*) and (*b*). Yet, despite Quine's arguments, it is my contention that he does not give us any inkling as to how this could be done.

References

Carnap, Rudolf 1955: "Meaning and Synonymy in Natural Languages". *Philosophical Studies* 6, 33-47.

Chomsky, Noam 1959: "A Review of B. F. Skinner's *Verbal Behavior*". *Language* 35, 26-58.

— 1969: "Quine's Empirical Assumptions". In: Jaakoo Hintikka & Donald Davidson (eds.), *Words and Objections*. Dordrecht: Reidel, 53–68.

Harman, Gilbert 1969: "An Introduction to 'Translation and Meaning' Chapter Two of *Word and Object*." In: Jaakoo Hintikka & Donald Davidson (eds.), *Words and Objections*. Dordrecht: Reidel, 14–26.

Mates, Benson 1951: "Analytic Sentences". *The Philosophical Review* 60, 525–534.

Montague, Richard 1973. "The Proper Treatment of Quantification in Ordinary English". In: Jaakoo Hintikka, Julius Moravcsik, and Patrick Suppes (eds.), *Approaches to Natural Language*. Dordrecht: Reidel, 221–242.

Quine, Willard van Orman 1951: "Two Dogmas of Empiricism". *The Philosophical Review* 60, 20–46.

— 1960: *Word and Object*. Cambridge, MA: MIT Press.

— 1969: "Reply to Harman". In: Jaakoo Hintikka & Donald Davidson (eds.), *Words and Objections*. Dordrecht: 295–297.

— 1987: "Indeterminacy of Translation Again". *The Journal of Philosophy* 84, 5–10.

Grazer Philosophische Studien
89 (2014), 161–183.

A TENSION IN QUINE'S NATURALISTIC ONTOLOGY OF SEMANTICS

Dirk GREIMANN
Universidade Federal Fluminense

Summary

According to Quine's naturalistic ontology of semantics, the semantic facts are limited to the facts about verbal behavior: if two speakers have the same speech dispositions, there can be no semantic differences between them. The aim of this paper is to show that Quine's ontology of semantics is incompatible with his principle of ontological commitment, broadly construed. The argument is that his ontology "presupposes" in a certain sense semantic facts whose existence is explicitly denied by it: when we confine the facts of semantics to facts about verbal behavior, our language undergoes a drastic loss of expressive power that implies that this theory cannot be formulated any more.

Introduction

Quine's ontological views are conditioned in large part by two general principles for the ontological recognition of entities. The first is the principle of ontological economy, which says that we must repudiate a given sort of entities if their recognition is not necessary ("entities must not be multiplied without necessity"). This principle is based on the more general principle of simplicity, according to which we must avoid any kind of unnecessary complexity in science. Applied to ontology, this implies that we must not make our ontology richer than necessary. The second principle is the principle of ontological commitment. It demands that we make our ontology as rich as necessary to account for the ontological presuppositions (or "implications") of the theories we accept. This principle derives from the more general principle of consistency, according to which we must not make claims or presuppositions that contradict one another.

In his ontology of semantics, Quine uses an additional and more special principle according to which the facts of semantics are exhausted by the facts of verbal behavior. Semantic questions whose answer is not determined by the facts about verbal behavior, are regarded by him as questions for which there is no "fact of the matter". Such questions are underdetermined not only by the empirical data, but also by the objective facts. In his view, there is, for instance, no fact of the matter deciding what the reference of a given term is, or whether two given terms are synonymous, or whether a given declarative sentence is analytic or synthetic, and so on. For this reason, he regards such distinctions as distinctions without an objective difference. They do not have any foundation in reality.

The rationale behind Quine's naturalistic approach is his effort to overcome the "myth of a museum" in mentalistic semantics according to which the semantic structure of our language is constituted by the correlation of ideas (as exhibits) with words (as labels) in our head (as a kind of mental museum). Quine does not deny that there are ideas in our heads; rather, he denies that they are part of the semantic structure of language. According to his naturalistic approach, this structure is not constituted in our heads, but in public, by the correlation of sentences with publicly observable verbal behavior. This means, technically speaking, that the semantic structure of our language is confined to the "stimulus meanings" of sentences. If two speakers have the same dispositions to affirm or deny given sentences in given publicly observable circumstances, then it does not make sense to imagine semantic differences between them.

In what follows, I shall argue that Quine's austere ontology of semantics offends against the principle of ontological commitment, broadly construed. This ontology is incoherent because it denies the existence of semantic facts that are in a certain sense "presupposed" by it. The problem is not that the existence of these facts is a condition for the truth of this theory, but a condition for its successful formulation and communication. When we reduce the semantic structure of our language to the correlation of sentences with stimulus meanings, then our language undergoes a drastic loss of expressive power that implies that we cannot formulate our theory any more. Hence, by asserting the theory, we are presupposing the existence of semantic structures that are denied by it. This is the incoherence that leads to a conflict with the principle of ontological commitment, as I shall try to show.

As far as I can see, this argument is not completely new, but it has never been spelled out in detail.[1] In section 1, I shall briefly recapitulate Quine's

1. See, for instance, Searle (1987, 131).

theory of ontological commitment. The task of section 2 is to explain the broader version of the principle of ontological commitment. Finally, in section 3, this criterion is applied to Quine's ontology of semantics in order to show that this theory is committed to recognize semantic structures whose existence it explicitly denies.

1. *Quine's theory of ontological commitment*

The task of a theory of ontological commitment is to specify which sorts of entities are presupposed by science. The theory sketched by Quine consists of four components: [i] an explication of the notion of ontological commitment; [ii] a criterion for deciding what the ontological commitments of a given theory are; [iii] a methodology of ontological reduction whose task is to eliminate ontological commitments that are not essential; and [iv] an application of this methodology to science telling us what the ontological commitments of science actually are.

[i] Quine never attempted to define or to give a detailed explication of his notion of "ontological commitment". He rather introduces this notion by means of examples like "we commit ourselves to an ontology containing centaurs when we say there are centaurs" (1948, 8). Nevertheless, from his examples it becomes clear that the kind of commitment he has in mind is basically the commitment to preserve coherence between our ontology, on the one hand, and the theories we accept in science, on the other. When we want to include a given theory T into our overall system of the world, we commonly have to settle an ontological bill for this: we must recognize, in our ontology, the entities presupposed by T. The obvious reason is that it is incoherent to deny the existence of the entities that are presupposed by the truth of the theories we accept. Suppose, for instance, that T asserts the sentence 'Snow is white'. Since the existence of snow is a condition of the truth of this sentence, T presupposes (or "implies") the existence of snow, even when in T the existence of snow is not explicitly asserted. Consequently, T becomes inconsistent when in T the existence of snow is denied. The kind of inconsistency involved in this case is semantic inconsistency: a conflict between the truth conditions of the sentences 'Snow is white' and 'Snow does not exist'.

An ontological commitment in Quine's sense is thus a *norm of coherence* for the construction of our ontology within our overall system of the world: when we take a theory T as true, and when the truth of T presupposes the existence of entities of the sort S, then, to be coherent, we must also take the ontological claim that the entities of the sort S exist as true.

To explain Quine's notion of ontological commitment more closely, it is important to take into consideration that his ontological development falls into two phases: the period before and the period after "Ontological Relativity" (1968). In the first phase, he assumes that theories have determinate ontological commitments. But, in "Ontological Relativity", he makes a profound revision of his ontological views that is due to the thesis of the indeterminacy of reference, according to which there are no facts determining the reference of our terms (cf. Quine 1968, 50). This thesis implies that there are also no facts determining what the ontological presuppositions of our theories are. To fix the intended interpretation of our terms, Quine relativizes reference to manuals of translation. Such a relativization is supposed to fix also the ontological presuppositions of the sentences occurring in the theory formulation. The notion of ontological commitment must hence be relativized in the same way as the notion of reference. I shall ignore this complication for the time being.

[ii] Generally speaking, a criterion of ontological commitment is any informative answer to the question which entities are presupposed by science. The task of Quine's criterion is to specify the entities that are presupposed by the truth of a theory and its sentences, respectively. Consider the sentence

(1) Fido is a dog.

There are basically two approaches to answer the question which entities are presupposed by the truth of (1), a platonist and a nominalist one. According to the former, (1) contains two referential expressions, namely, the singular term 'Fido', which refers to an object, and the general term 'is a dog', which refers to a property (or "universal"). The sentence is true if and only if there is an object denoted by 'Fido' and a property denoted by 'is a dog' and the object exemplifies the property. According to the nominalist approach, which Quine favors, general terms are not referential terms, but syncategorematic (non-referential) ones.[2] They do not denote

2. Cf. (Quine 1948, 10).

abstract objects, but they are satisfied or not satisfied by concrete objects. The truth conditions of (1) are that there is an object denoted by 'Fido' and that this object satisfies the general term 'is a dog'.

Strictly speaking, however, the sentences of natural language do not have clear and precise ontological presuppositions or implications, in Quine's view.[3] In order to determine the ontological commitments of a given theory, we must first translate its language into an ontologically transparent language. Quine thinks that the language satisfying this demand is the language of first order predicate logic with identity. On his construal, this language contains only one category of referential expressions, namely, first order variables. All other expressions, like the quantifiers, the truth functional connectives, the predicates and the functional signs, are considered as syncategorematic expressions. The singular terms are "explained away" by reducing them to definite descriptions in the way illustrated by the paraphrase of (1) as

There is one and only one x that fidos and this x is a dog.[4]

This step shifts the burden of reference from the singular term 'Fido' to the variable 'x'. Quine's criterion of ontological commitment accordingly reads: "To be assumed as an entity is, purely and simply, to be reckoned as the value of a variable" (Quine 1948, 13).

[iii] Roughly speaking, an "ontological reduction" is a translation of a given class of sentences whose truth appears to presuppose the existence of a certain sort of entities into sentences that do not have this presupposition. Its task is to show that the assertion of sentences of the first class does not really commit us to accept the entities in question.

Quine's methodology of ontological reduction consists mainly of two methods: contextual definition (Bentham's "paraphrase") and the method of proxy functions.[5] Suppose, for instance, that in a theory T the sentence 'The height of the Eiffel Tower = 324 m' is asserted. Since the singular term '324 m' occurs syntactically in a referential position, it seems that T is committed to recognize the existence of measures like 324 m. We can, however, show by means of the following contextual definition

3. Cf. (Quine 1953, 107).
4. Cf. (Quine 1948, 8).
5. Cf. Quine (1981, 14, 16ff.) and Quine (1981b, 68 ff). For a detailed reconstruction and critical evaluation of these methods, see Greimann (2009).

that all sentences about meters can be translated into sentences about pure numbers:

The height of x = y m if and only if the height of x in meter = y.[6]

Consequently, the assertion of 'The height of the Eiffel Tower = 324 m' does not really commit us to include such things like 324 m into our ontology. This commitment is not essential because it can be eliminated by means of a contextual definition.

Quine's second method of ontological reduction is the method of proxy functions.[7] It is designed to translate all sentences speaking about one sort of objects into sentences speaking about another sort of objects, where the proxy function maps the former onto the latter. Let f be a function mapping physical objects onto the space-time regions occupied by them. Then the open sentence 'x is a P' of the source language can be translated as 'x is the f of a P' into the target language. The sentence 'Fido is a dog', for instance, may be translated as

The space-time region of Fido is the space-time region of a dog.

On the intended reading, 'Fido' and 'dog' are syncategorematic expressions, not referential ones. The sentence says that the fidoing space-time region is a dogging space-time region. It speaks about the space-time region of Fido, not about Fido.

[iv] In "Things and Their Place in Theories" (1981), Quine aims to show that by means of the method of proxy function the domain of the language of science can be reduced to the domain of pure set theory. However, despite the overwhelming ontological economy that is gained in this way, he rejects this reduction for reasons of his "naturalism" and his "robust realism" with regard to physical objects.[8] The ontology Quine finally accepts is basically a moderate physicalist ontology that recognizes physical objects and also extensional abstract objects like sets. The latter are recognized by Quine because of the ontological commitment of physics and mathematics to abstract objects. This purely extensional ontology is considered by Quine to answer to all ontological commitments of science. Intensional

6. Cf. Quine (1976, 212f.) and Quine (1981, 14).
7. Cf. Quine (1968, 55ff.) and Quine (1981, 19f.).
8. Cf. Quine (1981, 21).

entities like properties and propositions are rejected in his ontology. They are considered as *entia non grata*, because their identity conditions are not sufficiently clear.[9]

2. *Ontological commitment and performative coherence*

In its most general form, the principle of ontological commitment says that a theory is committed to recognize those entities whose existence is presupposed by it. The version advocated by Quine is a more restricted principle, which reads: a theory is committed to recognize those entities whose existence is a condition of the *truth* of the sentences affirmed in it. On the intended reading, this principle specifies completely the necessary and sufficient conditions for being under the obligation to recognize given entities. It says that a theory is committed to recognize those *and only those* entities whose existence is a condition of the truth of the sentences affirmed in it.[10]

However, with regard to its completeness, Quine's principle is questionable, for at least two reasons. First, it does not account for the ontological commitments of normative theories like Kantian ethics. The categorical imperative is a non-descriptive sentence of the imperative form "Act in such a way that p!" This sentence is neither true nor false, but either valid or invalid. Nevertheless, Kantian ethics is not ontologically innocent, because the validity of a categorical norm presupposes the existence of absolute values.[11] When we incorporate Kantian ethics into our overall system of the world, we must hence extend our ontology in a significant way: we are committed to recognize the existence of a very queer sort of intrinsically prescriptive entities that provide categorical norms with their unconditional validity.[12]

Second, and more importantly, Quine's approach does not account for the ontological commitments deriving from the success conditions of the speech acts we make in scientific debates. This is the main difference between Quine's and Frege's ontology of semantics. Thus, Frege justifies

9. See, for instance, Quine (1995, 93).

10. Compare the following typical formulation of Quine's criterion of ontological commitment: "an entity is presupposed by a theory if and only if it is needed among the values of the bound variables in order to make the statements affirmed in the theory true" (1953, 108).

11. This is stressed by Kant (1785, 428f.) himself.

12. I am following Mackie (1977, 38) here.

the recognition of "objective" thoughts in his ontology by arguing that the existence of such entities is a condition for the successful communication in science. He writes:

> A thought does not belong specially to the person who thinks it, as does an idea to the person who has it: everyone who grasps it encounters it in the same way, as the same thought. Otherwise two people would never attach the same thought to the same sentence, but each would have his own thought; and if, say, one person put $2 \cdot 2 = 4$ forward as true whilst another denied it, there would be no contradiction, because what was asserted by one would be different from what was rejected by the other. It would be quite impossible for the assertions of different people to contradict one another, for a contradiction occurs only when it is the very same thought that one person is asserting to be true and another to be false. So a dispute about the truth of something would be futile. There would simply be no common ground to fight on; each thought would be enclosed in its own private world and a contradiction between the thoughts of different people would be like a war between ourselves and the inhabitants of Mars. Nor must we say that one person might communicate his thought to another and a conflict would then flare up in the latter's private world. It would be quite impossible for a thought to be so communicated that it should pass out of the private world of one person into that of another. (Frege 1897, 145; 1997, 233f.)

Obviously, this argument applies also to such speech acts as to answer a question and to discuss a hypothesis. Thus, in order to answer a question of a speaker A, the speaker B must determine the truth value of the thought expressed in the language of A. But, given the "idealist" interpretation of language, according to which thoughts are private, psychological entities, B cannot express the same thought in his language. Hence, he cannot answer any question asked by A.

According to Frege, a second success condition for communication in science is the existence of a common universe of discourse for all speakers of the scientific community. He criticizes the idealist logicians on the ground that their semantic interpretation of the language of science implies a certain kind of semantic solipsism. For, according to the idealistic interpretation, science speaks about mental representations. Since every individual has its own mental representations, the language of each speaker has his own universe of discourse. There is no common domain of objects to which all speakers refer. As a consequence, proper names like 'the Moon' must be treated as indexical terms whose reference depends on who is speaking. In the language of speaker A, this name refers to mental representations in the mind of A, and, in the language of speaker B to mental representations in the mind of B. This conception is criticized by Frege on the ground that it implies a kind of *linguistic solipsism* that undermines the possibility of successful communication in science:

> Thus everything leads into idealism and with perfect logical consistency into solipsism. If everyone designated something different by the name 'Moon', namely, one of his ideas, […], then admittedly the psychological way of looking at things would be justified; but a dispute about the properties of the Moon would be pointless: one person could quite well assert of his Moon the opposite of what another person, with equal right, said of his. If we could grasp nothing but what is in ourselves, then a [genuine] conflict of opinions, a reciprocity of understanding, would be impossible, since there would be no common ground, and no idea in the psychological sense can be such a ground. (Frege 1893, XIX; 1997, 206)

Given the idealist interpretation of scientific language, every speaker is a solipsist in the sense that his world, considered as the domain of his language, contains only himself and the contents of his consciousness. A consequence of this approach is, again, that it is impossible to contradict successfully the opinion of another speaker. This time, the problem is not that two different speakers A and B cannot express the same content, but, that they cannot refer to the same objects. The "common ground" of which Frege speaks is the realm of reference, and not the realm of sense. Suppose that A affirms that the object x has the property P. B wants to contradict this opinion. A success condition of this act is that B is able to refer to the same object x. *Ex hypothesi*, this is impossible. Consequently, B cannot contradict any of A's opinions.

Frege is certainly right that, by participating in a scientific debate, we are presupposing the existence of certain semantic structures that enable us to perform successfully such basic speech acts as making an assertion, asking a question, contradicting the opinion of an opponent, and so on. It is, consequently, incoherent when we deny in science the existence of those semantic structures that must be presupposed in order to be able to successfully perform such speech acts. This time, the incoherence is not a semantic, but a performative one. Assume that the existence of Fregean thoughts is a success condition for making assertions. Let T be a theory in which the existence of Fregean thoughts is denied. Then the assertion of T is performatively incoherent, because the truth of T implies that the conditions of the successful assertion of T are not fulfilled.[13] From this we may derive the following additional criterion of ontological commitment:

(AC) A theory is committed to acknowledge those semantic structures whose existence is a condition for the successful performance of the speech acts that are needed in order to communicate and to defend that theory.

Note that the ontological recognition of the semantic structures presupposed by the assertion of T is *justified in exactly the same degree* as the recognition of the entities whose existence is a condition of the truth of T. In both cases, the rationale behind the acknowledgement is the striving for consistency. The only difference is that in the case of the acknowledgement of the semantic structures the kind of consistency involved is performative consistency.

An important success condition for assertions is that the expressive power of the language we use is sufficient to express the intended contents of our assertions. Suppose, for instance, that we wish to assert a theory T whose objects are the natural numbers. To formulate T, we need a language whose semantic structure allows us to refer to the natural numbers and to describe their properties and relations. The semantic structure of first order predicate logic is not rich enough to formulate T. We need a language with a richer semantic structure as, for instance, the language of set theory or the language of second order predicate logic. When we assert T, we are hence committed to acknowledge the semantic structures of such richer languages. In particular, it would be incoherent to assert T and to deny, simultaneously, that there are any semantic structures that are not contained in the semantic structure of the language of first order predicate logic.

Generalizing this example, we can say that, in our ontology of semantics, we must acknowledge the semantic structures that are needed to provide our scientific language with the expressive power that is necessary to formulate our theories adequately. When we deny the existence of a given semantic structure, we normally have to pay a semantic price for this: we

13. This kind of incoherence is also illustrated by the Moorean paradox 'Snow is white, but I do not believe this'. Although there is no semantic conflict between 'Snow is white' and 'I do not believe that snow is white', because both sentences can be simultaneously true, there is a performative conflict between them, because they cannot be simultaneously asserted: for, by asserting 'Snow is white', the speaker expresses that he believes this.

must accept a corresponding loss of the expressive power of our language. The critical point is reached when the expressive power of our language becomes insufficient to formulate the theories we accept. In this case, we commit a performative inconsistency because we reject semantic structures that supply our language with the expressive power that is necessary to formulate the theories we accept.

3. *The performative incoherence of Quine's ontology of semantics*

It can be shown that Quine's anti-realistic ontology of semantics offends against our additional criterion of ontological commitment (AC). The truth of this theory implies a drastic reduction of the expressive power of our scientific language that makes it impossible to formulate that theory. In the end, it leads even to the linguistic solipsism criticized by Frege. Let me explain.

According to Frege's internalist approach, the semantic facts consist basically of the correlation of sentences with thoughts in the speaker's head. The meaning of a sentence is the thought that the speaker expresses by means of it. According to Quine's externalist approach, on the other hand, language is a public institution. The semantic structure of language is not determined by the unobservable correlation of sentences with thoughts in the speaker's head, but by the publicly observable correlation of sentences with verbal behavior. In "Ontological Relativity" (1968), he criticizes the internalist approach as follows:

> Uncritical semantics is the myth of a museum in which the exhibits are meanings and the words are labels. To switch languages is to change the labels. Now the naturalist's primary objection to this view is not an objection to meanings on account of their being mental entities, though that could be objection enough. The primary objection persists even if we take the labeled exhibits not as mental ideas but as Platonic ideas or even as the denoted concrete objects. Semantics is vitiated by a pernicious mentalism as long as we regard a man's semantics as somehow determinate in his mind beyond what might be implicit in his dispositions to overt behavior. It is the very facts about meaning, not the entities meant, that must be construed in terms of behavior. (Quine 1968, 27)

Unfortunately, Quine primary objection against mentalism is a mere reformulation of his naturalistic approach. The conclusion at which he wants to arrive is the naturalistic thesis that the semantic structure of our language ("the facts about meaning") is constituted exclusively by our verbal dispositions. If a given semantic property or relation is not determined by our verbal dispositions, then there is no fact of the matter determining whether or not given expressions have this property or stand in this relation. It is this ontological thesis which is presupposed by his arguments for the indeterminacy of translation, meaning, synonymy and reference. It should therefore be regarded as the core thesis of Quine's naturalistic conception of language.[14]

Technically speaking, this thesis is a *supervenience* principle saying that semantic properties and relations supervene on behavioral properties and relations in the sense that there can be no semantic difference between two languages when there is no behavioral difference between the speakers. Typical formulations of this principle are:[15]

> [...] if two speakers match in all dispositions to verbal behavior there is no sense in imagining semantic differences between them. (Quine 1960, 79)

> There is nothing in linguistic meaning [...] beyond what is to be gleaned from overt behavior in observable circumstances. In order to exhibit these limitations, I propounded the thought experiment of radical translation. (Quine 1987, 5)

> It is just that the factuality [of semantics] is limited to the verbal dispositions themselves, however elegantly or clumsily codified. Such, for me, are the facts of semantics. (Quine 1986, p. 155)

To illustrate this principle, suppose that an extraterrestrial lands on Earth with whom we can communicate perfectly well in English. His verbal behavior is indistinguishable from the verbal behavior of a native speaker of English. Still, the thoughts (beliefs) he expresses by uttering English sentences are radically different from the thoughts we usually express, and the objects to which he refers by using English terms are different from the objects to which we normally refer. Does the extraterrestrial speak English? Does he understand English? From Frege's point of view, both questions must be answered in the negative, because the correlation of sentences with the thoughts they express is an integral part of the semantic

14. This point is overlooked by Foellesdal (1990), (2011) and also by Gibson (2004). They construe Quine's behaviorism primarily as an epistemological thesis saying that the only evidence we can build on in our study of language is empirical evidence.

15. See also Quine (1968, 29) and Quine (1990, 110).

structure of English. According to Quine's supervenience principle, on the other hand, the extraterrestrial does speak English. Since his verbal behavior is identical to ours, there can be no semantic difference between his language and ours.

Surprisingly, the supervenience principle is a *dogma of Quine's naturalism*, that is, a central assumption that is taken more or less for granted, although it is in need of a thorough justification. There are only very sketchy arguments with which Quine underpins this important principle. In some writings, he seems to derive the principle from his "naturalism", considered as the methodology to treat language and the mind as objects of empirical science. Here are two examples:

> Let us […] recognize that the semantical study of language is worth pursuing with all the scruples of the natural scientist. We must study language as a system of dispositions to verbal behaviour, and not just surface listlessly to the Sargasso Sea of mentalism. (Quine 1975, 91)

> Philosophically I am bound to Dewey by the naturalism that dominated his last three decades. With Dewey I hold that knowledge, mind, and meaning are part of the same world that they have to do with, and that they are to be studied in the same empirical spirit that animates natural science. There is no place for a prior philosophy. When a naturalistic philosopher addresses himself to the philosophy of mind, he is apt to talk of language. Meanings are, first and foremost, meanings of language. Language is a social art which we all acquire on the evidence solely of other people's overt behavior under publicly recognizable circumstances. Meanings, therefore, those very models of mental entities, end up as grist for the behaviorist's mill. (Quine 1968, 26)

The "scruples" Quine has in mind seems to be the epistemic concern that the study of language must be restricted to empirical investigations that do not go beyond the empirical evidence. This argument, however, is not available to him, because, according to his thesis of underdetermination, the study of physical objects in natural science also goes beyond the empirical evidence. Moreover, in the speculative parts of his theory of language learning, Quine gives mentalistic explanations of how language is learned that are very remote from empirical evidence.[16] Thus, he postulates an innate instinct of "body-mindedness" (1974, §15) to explain how the

16. Especially in *Roots of Reference* (1974). For this reason, some authors assume that, during the phase of *Roots of Reference*, Quine abandoned the behaviorism he adopted in *Word and Object* (1960). See, for instance, Parsons (1990).

child learns the individuation of bodies and the reference to such entities. This procedure is also incompatible with the confinement of the study of language to strictly empirical investigations.

In *Word and Object*, Quine derives the supervenience principle from the uncontroversial claim that the empirical data on which a child relies when it learns his first language are limited to the observation of the verbal behavior of adult speakers. He writes:

> Language is a social art. In acquiring it we have to depend entirely on intersubjectively available cues as to what to say and when. Hence there is no justification for collating linguistic meanings, unless in terms of men's dispositions to respond overtly to socially observable stimulations. The effect of recognizing this limitation is that the enterprise of translation is found to be involved in a certain systematic indeterminacy [...] (1960, ix).

But this argument is a *non sequitur*, because there may be parts of our linguistic knowledge that are acquired by non-empirical mechanisms. Quine himself assumes that the reference to physical bodies is a linguistic skill that derives from an innate instinct of "bodymindedness" (1974, §15).

Finally, in some of his later writings, Quine infers the supervenience principle from a behaviorist standard of linguistic competence:

> In psychology one may or may not be a behaviorist, but in linguistics one has no choice. Each of us learns his language by observing other people's verbal behavior and having his own faltering verbal behavior observed and reinforced or corrected by others. We depend strictly on overt behavior in observable situations. As long as our command of our language fits all external checkpoints, where our utterance or our reaction to someone's utterance can be appraised in the light of some shared situation, so long all is well. Our mental life between checkpoints is indifferent to our rating as a master of the language. There is nothing in linguistic meaning, then, beyond what is to be gleaned from overt behavior in observable circumstances. (Quine 1987, 5)

It seems, however, that Quine is begging the question here. From the mentalistic point of view, linguistic competence is not merely a question of "fitting all external checkpoints", but also a question of "expressing the right thoughts". Our extraterrestrial does not count as a competent speaker of English, on the mentalist's standard.

Leaving Quine's arguments aside, the philosophically interesting question is how to decide, in a justified way, what the extension of the semantic facts is. Clearly, this question cannot be decided by means of an empirical investigation. Nor does it help to analyze the notion of language. The method suggesting itself is to apply our additional principle of ontological commitment: to determine the extension of the semantic facts, we must determine which semantic structures we presuppose in scientific communication. We can then show, negatively, that Quine's supervenience principle is unduly restrictive, because it implies that his own theory cannot be formulated in an adequate way. The argument is as follows.

To show that an alleged semantic structure does not supervene on the dispositional structure of language, Quine uses his thought experiments of the indeterminacy of meaning and reference. They aim to show that the semantic structure in question can be permuted without affecting the dispositional structure. Take, for instance, the correlation of terms with their extensions. The thought experiment of the indeterminacy of reference shows that this correlation can be permuted without affecting the dispositional structure of language. This result implies that it is possible that in the languages of two speakers the correlation of terms with extensions is not identical, although the speakers match in all dispositions to verbal behavior. Since this outcome contradicts the supervenience principle, the correlation of terms with extensions cannot be accepted in Quine's ontology of semantics: there are no facts determining what the extensions of terms are. The predicate 'x is the extension of the term t in L' must accordingly be considered as a scientifically non-respectable predicate that does not have an extension. To resolve this problem, Quine relativizes the extension of terms to manuals of translation.

In *Word and Object*, Quine explains more closely what the semantic facts are, in terms of the notion of stimulus meaning.[17] Formally, the stimulus meaning of a sentence is the ordered pair consisting of its positive and its negative stimulus meaning. The positive stimulus meaning is the class of stimuli that prompt the speaker to assent to the sentence when he is asked; the negative stimulus meaning is accordingly the class of stimuli that prompt him to dissent. The stimuli are considered by Quine as types, not as tokens.[18] It is hence perfectly possible that, in the languages of two different speakers, a sentence has exactly the same stimulus meaning. Thus, Frege's objection against the idealist interpretation of language that it

17. Cf. Quine (1960, chap. 2, especially 39).
18. Cf. Quine (1960, 34).

implies that meanings are private objects does not apply, *mutatis mutandi*, to Quine's behaviorist interpretation.

In Quine's view, the notion of stimulus meaning can be considered as an adequate explication of the meaning of observation sentences; their meaning consists in their stimulus meaning.[19] But the same does not also apply to theoretical sentences: since there are no or only very few stimuli prompting the speaker to assent or to dissent to such sentences, they have almost all the same stimulus meaning. Theoretical sentences have meaning only in an indirect way, *via* the implication of observation sentences. Since, however, a theoretical sentence does not imply any observation sentence individually, but only together with other theoretical and observational sentences, it does not have meaning individually, but only holistically, that is, as a part of a larger system of sentences.[20]

Finally, the words of which theoretical sentences are composed have meaning only in a doubly indirect way: their meaning consists in the contribution they make to the meaning of the sentences in which they occur.[21] A word has meaning only in the context of a sentence, and a sentence has meaning only in the context of a larger system of sentences.

The basic language-world relation is hence the correlation of observation sentences with stimulus meanings. These sentences have a *holophrastic* contact with the world, and not a compositional one, that is, they are connected with the world only as whole sentences, and not by the mediation of words. The entities in the world with which sentences are correlated are the stimuli that prompt either assent or dissent to these sentences. This does not mean, of course, that an observation sentence like 'This is a dog' speaks about stimuli. Nor do these sentences speak about objects like dogs. Rather, they do not speak about anything because they do not contain any referential expressions at all; they are ontologically neutral. Considered as an observation sentence, 'This is a dog' does not say of an object that it is a dog, but only that "it is dogging".

Since observation sentences are occasion sentences, their truth value depends on the context of utterance. The truth conditions of 'This is a dog', considered as an observation sentence, may be explained as follows:

---

19. Cf. Quine (1960, 42).

20. See, for instance, Quine (1968b, 79).

21. In Quine's words: "[…] the meanings of words are abstractions from the truth conditions of sentences that contain them" (1981b, 69).

> The sentence 'This is a dog' is true in the context C if and only if the pattern of stimuli that the speaker receives in C belongs to the positive stimulus meaning of this sentence.

Note that the presence of a dog in C or the existence of any dog is *not* a condition for the truth of the sentence in C. Its truth depends only on the triggering of the receptors of the speaker in C. As a consequence, observation sentences cannot be used to speak about any kind of objects. The reason is that the correlation of sentences with stimulus meaning is a very poor semantic structure that does not allow us to refer to objects. Since theoretical sentences receive their meaning only indirectly from the stimulus meanings of the observation sentences they imply, these sentences cannot be used to speak about any kind of objects, either. Theoretical sentences are connected with the world only indirectly, *via* the implication of observation sentences, which are in turn connected with the world via their stimulus meaning. Consequently, theoretical sentences and the words occurring in them are connected only with stimuli, but not with objects. In a language with a behaviorist interpretation, there are no referential terms at all.

Note that it does not help to relativize the notion of reference in the way envisaged by Quine in order to allow us to speak about objects in a language with a behaviorist interpretation. Insofar as there are facts determining the extension of the relativized predicate 'x is the extension of the term t in L relative to the manual M', the relativization may help us to construe a scientifically respectable notion of reference. But, since this maneuver does not enhance the expressive power of our language—the relativization does not add any semantic structure to our language—, it does not enable us to speak about objects. The problem is that the expressive power of a language with a behaviorist interpretation is fully captured by the stimulus meanings of its ontologically neutral observation sentences. Let S be the set of the observation sentences implied by a theory formulation F. Then everything that is expressed by F is also expressed by S. Since the sentences of S do not speak about any objects, neither do the sentences of F.

It would be misleading to say that in a language with a behaviorist interpretation the reference of terms is indeterminate. This suggests that such a language does contain referential terms, albeit referential terms whose reference is not fixed. In fact, however, such a language does not contain any referential terms at all. Consequently, its referential structure is completely determined, if only in a negative way.

It follows that Quine's ontology of semantics is performatively inconsistent. In order to formulate this theory, we must speak about such objects as speakers, sentences, verbal dispositions, and so on. On the other hand, the theory implies that the semantic structures that enable us to speak about such objects do not exist. Hence, by asserting the theory, we presuppose the semantic structures whose existence we deny. As Searle puts it: "*If the indeterminacy thesis were really true, we would not even be able to understand its formulation*" (1987, 131).

It can, moreover, be shown that the behaviorist interpretation of language collapses into Frege's semantic solipsism. We saw that, on the behaviorist interpretation, the truth of a sentence in a context of utterance depends exclusively on the patterns of stimuli that the speaker receives in this context. We can imagine a situation in which a speaker A affirms a sentence which a speaker B denies and both A and B are right. Such a situation is described by Davidson in his critique of Quine's "proximal" theory of meaning according to which the truth value of a sentence depends on the stimuli that the speaker receives:

> [...] let us imagine someone who, when a warthog trots by, has just the patterns of stimulations I have when there is a rabbit in view. Let us suppose the one-word sentence the warthog inspires him to assent to is 'Gavagai!' Going by stimulus meaning, I translate his 'Gavagai!' by my 'Lo, a rabbit' though I see only a warthog and no rabbit when he says and believes (according to the proximal theory) that there is a rabbit. The supposition that leads to this conclusion is not absurd; simply a rearranged sensorium. Mere astigmatism will yield examples, deafness others; little green man and women from Mars who locate objects by sonar, like bats, present a more extreme case, and brains in vats controlled by mad scientists can provide any world you or they please. (Davidson 1990, 74)

In the language of A, the truth value of 'Gavagai!' in a context C depends on the events taking place at A's sensory receptors in C, whereas, in the language of B, the truth value of the same sentence and in the same context depends on the events taking place at B's sensory receptors in C. Since the pattern of stimuli that the person with astigmatism receives when a warthog trots by belongs to the positive stimulus meaning of 'Gavagai!', this person is right to assent to 'Gavagai!' when a warthog trots by, although no rabbit is present.[22] At the same time, we are right to deny

22. Davidson (1990, 74) assumes—erroneously, I think—that the person with astigmatism is wrong when he asserts 'Gavagai' in the situation described.

'Gavagai!' (or its translation 'Lo, a rabbit!') in the same situation, because the pattern of stimuli we receive when a warthog trots by belongs to the negative stimulus meaning of 'Gavagai!'. It is, therefore, impossible in a language whose semantic interpretation is restricted to the assignment of stimulus meaning that a speaker A contradicts the opinion of a speaker B. The problem is the lack of a "common ground" in Frege's sense: the truth values of the sentences of such a language are not determined by a common public world, but by the solipsistic worlds of each individual speaker.

To overcome this difficulty, we must assume that the truth conditions of our sentences do not depend on the properties of the speaker – which of his sensory receptors are triggered –, but on the properties of the objects that these sentences typically are about. Thus, the truth conditions of 'Gavagai!' (and 'Rabbit!") are that a real rabbit is present in the context of utterance. Davidson calls this the "distal" theory of meaning and truth (cf. 1990, p. 73). Quine himself seems to adopt it in the following passage from *Word and Object*, where he explains more closely the public character of language:

> Linguistically, and hence conceptually, the things in sharpest focus are the things that are public enough to be talked of publicly, common and conspicuous enough to be talked of often, and near enough to sense to be quickly identified and learned by name; it is to these that words apply first and foremost. (Quine 1960, 1)

Some commentators assume that the distal approach fits better into Quine's naturalistic-behavioristic picture of language, because it does more justice to the public character of language.[23] In my own view, however, the distal view is incompatible with Quine's assumption that physical objects (including rabbits and other "physical bodies") are theoretical constructs whose ontological recognition goes beyond all possible observations. In contrast to Davidson, Quine does not consider physical objects as publicly observable entities. This becomes clear from the following passage from "Things and Their Places in Theories":

> [...] I see all objects as theoretical. This is a consequence of taking seriously the insight that I traced back from Bentham—namely, the semantic primacy of sentences. It is occasion sentences, not terms, that are to be seen as conditioned to stimulations. [...] The objects, or values of variables, serve merely

23. See, for instance, Føllesdal (2011, 277).

> as indices along the way, and we may permute or supplant them as we please as long as the sentence-to-sentence structure is preserved. (Quine 1981, 20)

Suppose, for instance, that our extraterrestrial does not have an innate instinct of body-mindedness, but of bundle-mindedness: when he looks at a rabbit, he does not "see" a body, but a mere bundle of qualities. Nevertheless, he speaks fluently English. This already shows that rabbits, construed as bodies, are not publicly observable entities, but theoretical constructs that we posit in order to systematize our observations. Physical objects are not discovered by us; they are invented by us.

Even when we accept the distal view, the supervenience principle still implies that we cannot assert such things as the existence of a rabbit in this room. To assert this, we must formulate a sentence whose truth depends on the existence of rabbits. To achieve that, we must fix the reference of 'rabbit' in such a way that it is satisfied only by rabbits. Given the supervenience principle, the only means we have to do that is our publicly observable verbal behavior. We must hence fix the intended reference of 'rabbit' by gestures and verbal utterances that amount to an ostensive definition of rabbits. To achieve such a definition, we must point to rabbits and stipulate that the indicated objects are "rabbits". As a matter of fact, however, mere ostension is incapable of fixing the intended reference. For, when we point to a rabbit, we are always pointing to very different things: to a body, for instance, and also to the space-time region occupied by a body and to a temporal segment of a body. Hence, our verbal behavior does not fix what the referent of 'rabbit' is supposed to be. This finally implies that we cannot formulate a sentence whose truth depends on the presence of a rabbit in this room.

It could be objected that, for Quine, a rabbit is not *per se* a rabbit, because there are many other sorts of objects that can play the role of rabbits in our theories.[24] Since certain space-time regions can play this role, although they are not physical objects, to be a rabbit does not necessarily imply to be a physical object. To be a rabbit (or a "version" of rabbits) is simply to be a possible value of the variables (an "index") in a theory about rabbits. Given this structuralist notion of physical objects, which is suggested by Quine's thesis of ontological relativity, it is indeed possible to express, in a language satisfying the supervenience principle, that there is a rabbit in this room (whatever a rabbit may be).[25] By asserting that there is

24. This reading is suggested by Hylton (2007, 317–323), (2004, 144f.), and (2000, 298f.).
25. See, for instance, Quine (1968) and (1981, 20).

a rabbit in this room, we are merely asserting that there is some version of rabbits in this room, without saying which version this is supposed to be. Hence, the truth of the sentence 'There is a rabbit in this room' depends on the existence of rabbits, although it leaves open which version of the rabbits (bodies or space time regions of a certain kind, etc.) exists.

However, although there are some places in Quine's work where he sympathizes with this structuralist kind of physicalism, he officially defends the "robust" kind of physicalism according to which physical objects *qua* physical objects are the prime reality. The ground elements of Quine's ontology are not just physical objects in the abstract sense that allows us to construe them also as space-time regions or sets, but physical objects in the concrete sense according to which any space-time region or set is not a physical object.[26] This is important for Quine because he does not want to regard all versions of physical objects as equally adequate models of reality.[27]

It is obvious that, to formulate robust physicalism, we must be able to formulate a sentence whose truth depends on the existence of physical objects in the robust sense. But this is not possible when our language satisfies the behaviorist supervenience principle. The problem is, again, that we cannot fix, by means of our verbal behavior, what the intended reference of 'physical object' is supposed to be, even when we presuppose that the truth conditions of our sentences do not depend on our surface irritations, but on the properties of the objects at which we typically point when we learn our first language.

The conclusion to be drawn is that Quine's ontology is incoherent in the performative sense. Its truth implies that it cannot be asserted in a successful, non-defective way. In order to account for the ontological commitments deriving from the conditions for successful communication in science, we must reject the behaviorist supervenience principle. To be sure, this step is not acceptable for Quine, because, without the supervenience principle, his indeterminacy theses lose their plausibility. For this reason, the supervenience principle may turn out to be the Achilles' heel of his philosophy.[28]

---

26. See Quine (1995, 41). For more details, see the reconstruction in Greimann (2000).

27. See Quine (1981, 21).

28. For more details on the status of Quine's behaviorism in his overall philosophical system, see Gibson (1982, especially xx, 205).

Davidson, Donald 1990: "Meaning, Truth and Evidence". In: Robert B. Barrett & Roger F. Gibson (eds.), *Perspectives on Quine.* Oxford: Blackwell, 68–79.

Føllesdal, Dagfinn 1990: "Indeterminacy and Mental States". In: Robert B. Barrett & Roger F. Gibson (eds.), *Perspectives on Quine.* Oxford: Blackwell, 98–109.

— 2011: "Developments in Quine's Behaviorism". *American Philosophical Quarterly* 48, 273–282.

Frege, Gottlob 1997: *The Frege Reader*. Edited and introduced by Michael Beaney, Oxford: Blackwell.

— 1893: *Grundgesetze der Arithmetik. Begriffsschriftlich abgeleitet*, Vol. 1. Jena: Hermann Pohle. Reprint: Darmstadt: Wissenschaftliche Buchgesellschaft, [2]1962.

— 1897: *Logik*. Posthumous writing. In: Gottlob Frege, *Nachgelassene Schriften und Wissenschaftlicher Briefwechsel*, Vol. 1. Edited by H. Hermes, F. Kambartel and F. Kaulbach, Hamburg: Meiner. Second and extended edition 1983, 137–63.

Gibson, Roger F. Jr. 1982: *The Philosophy of W.V. Quine. An Expository Essay*. Tampa: The University of South Florida Press.

— 2004: "Quine's Behaviorism cum Empiricism". In: Roger F. Gibson (ed.), *The Cambridge Companion to Quine*. Cambridge: Cambridge University Press, 181–199.

Greimann, Dirk 2000: "No Entity without Identity: A Reductionist Dogma?". *Grazer Philosophische Studien* 60, 13–29.

— 2009: "Contextual Definition and Ontological Commitment". *Australasian Journal of Philosophy* 87, 1–17.

Hylton, Peter 2000: "Reference, Ontological Relativity, and Realism". *Aristotelian Society Supplementary Volume* 74, 281–299.

— 2004: "Quine on Reference and Ontology". In: Roger F. Gibson (ed.), *The Cambridge Companion to Quine*. Cambridge: Cambridge University Press, 115–150.

— 2007: *Quine*. London and New York: Routledge.

Kant, Immanuel 1785: *Grundlegung zur Metaphysik der Sitten*. Akademie-Ausgabe Vol. IV, Berlin, 1968.

Mackie, John L. 1977: *Ethics: Inventing Right and Wrong*. New York: Penguin Books.

Parsons, Charles 1990: "Genetic Explanation in *The Roots of Reference*". In: Robert B. Barrett & Roger F. Gibson (eds.), *Perspectives on Quine.* Oxford: Blackwell, 273–290.

Quine, Willard Van Orman 1948: "On What There Is". In: Willard V.O. Quine, *From a Logical Point of View*. New York: Harper and Row. First published by Harvard University Press in 1953, with a second, revised edition in 1961, 1–19.
— 1953: "Logic and the Reification of Universals". In: Willard V.O. Quine, *From a Logical Point of View*. New York: Harper and Row. First published by Harvard University Press in 1953, with a second, revised edition in 1961, 102–129.
— 1960: *Word and Object*. Cambridge, MA: The MIT Press.
— 1968: "Ontological Relativity". In: Willard V.O. Quine, *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969, 26–68.
— 1968b: "Epistemology Naturalized". In: Willard V.O. Quine, *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969, 69–90.
— 1974: *The Roots of Reference*. La Salle, Illinois: Open Court.
— 1975: "Mind and Verbal Dispositions". In: Samuel Guttenplan (ed.), *Mind and Language*. Oxford: Clarendon Press, 83–95.
— 1976: *The Ways of Paradox and Other Essays*. Revised and enlarged edition. Cambridge, MA: Harvard University Press.
— 1981: "Things and Their Place in Theories". In: Willard V.O. Quine, *Theories and Things*. Cambridge, MA: Harvard University Press, 1–23.
— 1981b: "Five Milestones of Empiricism". In: Willard V.O. Quine, *Theories and Things*. Cambridge, MA: Harvard University Press, 67–72.
— 1987: "Indeterminacy of Translation again". *Journal of Philosophy* 84, 5–10.
— 1995: *From Stimulus to Science*. Cambridge, MA: Harvard University Press.
Searle, John R. 1987: "Indeterminacy, Empiricism, and the First Person". *Journal of Philosophy* 84, 123–146.

Grazer Philosophische Studien
89 (2014), 185–203.

IN DEFENSE OF QUINE'S OSTRICH NOMINALISM

Guido IMAGUIRE
Universidade Federal do Rio de Janeiro

Summary

My aim in this paper is to defend Quine's so-called "Ostrich Nominalism". This pejorative designation was introduced by Armstrong (1978), who accused Quine of not taking predicates with "ontological seriousness" and not offering any explanation of predication. However, I think Quine is entirely correct in this. In particular, I will show how to counter the Platonist One Over Many argument for the existence of universals. In doing this, I will go beyond merely offering an exegesis of Quine's views on ontology: I will also try to fill in some gaps in his original argumentation.

Introduction

In "On What There is" (1948), Quine proposed a very general principle designed to enable us to decide virtually all questions of existence, the principle of "ontological commitment". As I understand it, this principle states that our ontological claims must be coherent with the theories we accept. Our best overall theory of the world gives us many sentences we should hold to be true. In order to decide questions of existence, we have to examine the ontological import of these sentences. To this end, we must translate these sentences into the canonical notation of first order logic, which Quine regarded as the best logic we have. In this logic, there is a rule of existential generalization stating that from "a is F" we may infer "There is an x such that x is F". This rule implies that we cannot coherently affirm that *a* is *F* and in the next breath deny that there is an *F* or that *a* exists. When we affirm that *a* is *F*, we must necessarily conclude that the particular *a* exists. But, on the other hand, it is not necessary to conclude that *F*-ness exists. According to Quine there are red houses and red roses, but not redness.

There is an important point that we must remember when we employ this procedure to make ontological decisions: we must distinguish the merely apparent ontological commitments of a theory from the genuine ones. For example, a theory containing the sentence "There is the possibility that S" seems to commit us to the existence of possibilities, which are very strange kind of abstract entities that nominalists would not be happy to accept. Actually, however, this commitment is only apparent; it can be avoided by replacing this sentence with an ontologically modest paraphrase like "It may be true that S".

However, it would be wrong to conclude that this simple strategy enables us to eliminate all abstract entities from our ontology. We are committed, as Quine himself concedes, to accepting the existence of numbers. From a nominalist viewpoint, this concession could be considered a weakness. But, from a meta-metaphysical perspective, it makes Quine's position even stronger. This position enjoys the reputation of being "ontologically correct" in the sense of doing justice to its own ontological commitments: despite his nominalist tendencies, Quine grudgingly accepted a Platonist ontology of mathematics.

Nevertheless, with regard to the traditional Problem of Universals, which refers mainly to the question of the existence of properties, Quine holds to the nominalist position, arguing that *Fa* does not commit us to the existence of *F*-ness, but only to that of *a*. This straightforward solution to the Problem of Universals was called by Armstrong "austere" or "ostrich nominalism". Many theorists are opposed to this kind of nominalism, including even some nominalists. But it also has its defenders, such as Michael Devitt (1980), James van Cleve (1994) and Joseph Melia (2005).

In this paper, I defend the view that Quine's ostrich nominalism is the best answer nominalists can offer to the Platonist One Over Many argument in favor of the existence of universals. Some authors have argued that Quine left many gaps in his argumentation. I agree with them, but I think these gaps can be filled. This will be the main task of this paper. However, in order to fill the gaps, I make extensive use of resources that are alien to Quine's philosophy, such as, for instance, the notion of truthmaking. I do not claim that my defense of ostrich nominalism would be acceptable to Quine, but only that his approach can be defended in this way. I call this position "New Ostrich Nominalism".

The gaps I discuss derive from two criticisms of Quine's solution. The first goes back to Armstrong (1980), who claims that Quine does not take predicates with ontological seriousness:

> [Quine defended the] extraordinary doctrine that predicates involve no ontological commitment. In a statement of the form "Fa", he holds, the predicate "F" need not be taken with ontological seriousness. Quine gives the predicate what has been said to be the privilege of the harlot: power without responsibility. (Armstrong 1980, 104f.)

As we shall see, this criticism cannot be sustained. The second criticism has been made by many philosophers, including Alston (1958), Mellor and Oliver (1997, 13–16), and could be formulated as follows:

> Suppose that sentence P is a paraphrase of sentence Q, and P commits us to the existence of the entity E, while Q does not. Why should we say, in this case, that the commitment of P to E is only apparent? Could we not also say that the non-commitment of Q to E is only apparent, so that Q does commit us to E after all?

This is truly an interesting point. I will call this problem "the paraphrase commitment problem". A solution is not easy to find, and insofar as I can see, Quine never solved it. This is what I attempt to do here. In the first part of this paper, I more closely examine the Problem of Universals and, in particular, the specific problem posed for nominalism by the Platonic One Over Many argument. In the second part, I present a solution for this problem based on Quine's ostrich nominalism.

1. *The problem of universals and the Platonic One Over Many Argument*

In the literature, there is no agreement about the most appropriate formulation of the Problem of Universals. At least three formulations have been offered by contemporary metaphysicians:

F1 Are there universals?
F2 How can different particulars be identical in nature?
F3 In virtue of what is the particular *a* an *F*, if the *a* is an *F*?

These are related, but different questions. The first is a question about what exists. The second is a question about how to explain that a property can be shared by different particulars. The third, finally, is a question about ontological grounding. I will not discuss in detail which of them is the most appropriate formulation. One can assume, as Campbell (1990) does, that F2 does not really express a problem, but rather an argument for the existence of universals: particulars can be identical in nature precisely because there are universals that are simultaneously exemplified by them. As I see things, with F2 Armstrong tries to capture the essence of the original One Over Many argument, which derives the existence of universals from the fact of multiple instantiation. Ostrich nominalism is simply a straightforward rejoinder to this argument. Something similar could be said about F3. The question "In virtue of what is the particular *a* an *F*?" suggests the Platonic answer that *a* is an *F* in virtue of exemplifying *F*-ness. In what follows, I assume that the Problem of Universals is basically a problem about the existence of F-ness, *i.e.*, F1 is the basic formulation. I do not think that the ostrich nominalist has a solution to all problems connected with the existence of universals. In particular, he does not have a solution for problems posed by second order predications. It might be that, in the end, the ostrich nominalist will have to accept the existence of universals, but in this case, the reason is not that there are many *F*s, but rather that quantification over *F* is unavoidable.

To assume that the Problem of Universals is a problem about the existence of universals has an important consequence for our approach. Rodriguez-Pereyra (2000) argued—correctly, I think—that the strategy of solving the problem by determining the truthmakers of our sentences and the alternative strategy of analyzing our ontological commitments lead to very different approaches of philosophical analysis. The procedure based on truthmaking goes from ontology to language: something makes a sentence true, whereas, the procedure based on ontological commitment runs from language to ontology: a sentence commits us to the existence of something. Which approach should we prefer when we deal with the Problem of Universals? Rodriguez-Pereira argues for the *truthmaker* approach and the ostrich nominalist for the *commitment* approach.

I think that when we assume that the Problem of Universals is a problem of existence, we must prefer the ontological commitment approach for a very simple reason. Truthmakers cannot give us what we are looking for. Truthmaking is a relation of necessitation. A given entity E makes the sentence S true if and only if E necessitates the truth of S. Thus, the existence of E entails the truth of S. But the important point is this: the truth of S does not entail the existence of E, because different entities can make S true. Take, *e.g.*, the following three different truthmakers of the sentence "Fa": (1) the universal *F*-ness being instantiated by the particular *a*, (2) the single state of affairs *Fa* and (3) the trope *F*-ness-of-*a*. It is clear that the existence of each of them alone would be enough to make the sentence “Fa” true. But then, given that our only basic datum is the truth of “Fa”, we cannot decide whether we should accept universals, states of affairs, tropes or all of them in our ontology, because it could be that all three exist, or only two, or only one of them—in any case the sentence remains true. When we want to decide on the existence of a given sort of entity, and the only data we have is the truth of a collection of sentences, we must work from language to ontology, and not *vice versa*. Consequently, the determination of ontological commitments and not the determination of truthmakers is the appropriate method that we should choose. But this does not imply that truthmakers are dispensable in our analysis. On the contrary, they play an important role in my proposal.

A second methodological assumption I make is that the metaphysical problems discussed here are not pseudo-problems. Much has changed since the days when we could simply reject metaphysical problems in general as language mistakes. Few contemporary philosophers reject metaphysical problems as mere pseudo-problems. And even those who tend to reject traditional metaphysical problems, like Hirsch (2009), accept the nominalist versus Platonist debate as a genuine philosophical problem. Quine, who was no friend of old-fashioned metaphysical inquiry, clearly accepted the genuineness and the substantiality of the debate over the existence of properties.

Now, if this debate expresses a genuine problem, then we should adopt the “principle of substantiality”, as I would like to call it. This principle states that no solution should solve a problem in a trivial, unsubstantial manner. Any solution that makes Platonism or nominalism trivially true or false should be rejected as inadequate. For example, a solution that conceives the problem of universals in terms of F1 and defines existence as location in space and time makes Platonism trivially false and consequently must be rejected. Indeed, in the passage quoted above, Armstrong accuses Quine of trivializing the problem in favor of nominalism. I do not think that this criticism is correct. Of course, we should not give predicates power without responsibility, ignoring their ontological input. But, as we will see, this is not Quine’s suggestion. On the contrary, the ostrich nominalist avoids trivialization in his own way.

Before we examine the solution of the ostrich nominalist in detail, let us see whether the two realist alternatives are really more substantial. The first was proposed by Russell, the second by Armstrong himself. According to Russell’s position, at least as explained by Donagan (1963), only the referents of non-redundant and non-formal predicates exist. Thus, the use of the predicate "is identical to" does not commit us to recognizing the existence of the entity *identity,* because it is a formal term. And the use of the predicate "human" does not commit us to recognizing the existence of universal humanity, because it is not primitive (it can be reduced to, say, "rational animal"). But since "rational" is a primitive predicate (let us suppose), we are compelled to accept the existence of the universal *rationality*. And, as strange as it may seem, even if no particular were rational, rationality would nevertheless exist, for in this case the sentence "a is not rational" would be true for any *a*. Indeed, anyone who uses this criterion is very near to an unsubstantial form of Platonism. This criterion for deciding if there is a corresponding universal is too easily fulfilled. We can hardly deny that we must use some irreducible predicates, and thus it becomes unavoidable for us to accept the existence of universals.

Armstrong proposes a different strategy for avoiding unsubstantial Platonism. In his view, only predicates that "carve the great beast of reality at the joints" commit us to universals. Only the predicates offered by the natural sciences, particularly by physics, express genuine properties. But again Platonism becomes unavoidable. It is evident that every theory has primitive predicates, and that we need predicates to formulate the most fundamental scientific laws. Nobody would dare to defend the contrary position.

But Quine also adopted a substantial, non-trivial position. We should not simply assume that there are properties just because there are predicates. Instead, an additional, more substantial criterion is needed. For Quine, as we know, the quantificational criterion is the decisive one. We must accept the properties expressed by the predicates we quantify over in our best overall theory—insofar as it is not possible to give a first order paraphrase. Indeed, and this is my point here, compared with Russell and Armstrong, Quine's suggestion is the most substantial one. It is far from trivial whether we can formulate every sentence of our overall theory in first order logic. It is also far from trivial whether a sentence that quantifies over predicates has a first order paraphrase. Since Quine's day, an extensive discussion on paraphrases of second order sentences has shown this in a very striking way. Thus, we can hardly regard Quine's position as a kind of unsubstantial nominalism.

Nevertheless, a far-reaching problem remains for the ostrich nominalist. Some have argued against Quine´s suggestion to "paraphrase ontological commitment away". Suppose that sentence P is a paraphrase of sentence Q and that P, but not Q, commits us to the existence of the entity E. Why should we conclude, as Quine suggests, that the commitment of P to E is only apparent? Since being a paraphrase of something is a symmetrical relation, one could also conclude that the non-commitment of Q to E is only apparent. This "Paraphrase Problem", as I called it at the beginning, is the main topic of the rest of this paper.

2. *The solution of Ostrich Nominalism*

Armstrong (1978) conceived the problem of universals basically in terms of formulation F2: How can different particulars be identical in nature? But, as Oliver (1996, 49-50) noted, Armstrong tried to answer this question by offering an account of the facts stated by six sentences that are not equivalent, namely:

(1) *a* and *b* are of the same type/ have a common property
(2) *a* and *b* are both *F*
(3) *a* and *b* have a common property, *F*
(4) *a* has a property
(5) *a* is *F*
(6) *a* has the property *F*

Oliver is right in holding these to be different sentences with different ontological commitments. For my purposes, I will add one more sentence:

(7) *b* is *F*

This is our first example, which we will call "EX1". These sentences will help us understand how the ostrich nominalist sees the Platonist "One Over Many" argument. According to this argument, if it is the case that *a* is *F,* and also that *b* is *F*, then (5) and (7) are true. Thus, it is the case that *a* and *b* are both *F*, and so (2) is true. But if (2) is true, (3) must also be true: *a* and *b* have a common property *F*. Finally, from (3) the Platonist derives the conclusion: *there is* something that *a* and *b* have in common. This something is the property *F*. Thus, there is *F*-ness. Note that in this argument we are just applying Quine's own quantificational criterion. So far, the ostrich and the Platonist agree at least about the criterion. Now, how could we counter this conclusion?

2.1 *First step: minimalism*

The point I want to clarify now is very simple. Let us suppose that we accept a given theory as true. A theory is basically a set of sentences. Thus, we assume that sentences S_1, S_2, S_3..., S_n, are all true. And suppose that sentence S_1 commits us to the existence of entity E_1, sentence S_2 to the existence of entity E_2, sentence S_3 to the existence of entity E_3, and so on. Should we conclude that this theory commits us to the existence of all entities E_1–E_n? The ostrich nominalist regards this as *not* unavoidable. Quine never explicitly proposed the strategy I present here. But I think I have found a way to make sense of his claim that, "One may admit that there are red houses, roses, and sunsets, but deny, except as a popular and misleading manner of speaking, that they have anything in common" (1948, 10).

This is our situation in EX1. We have sentences we accept as true. Now, the point is: It must be clear to everyone that the set of sentences (1)–(7) is in an informational sense redundant. We have a situation (that *a* and *b* are *F*s) that is described using "too many" sentences. I call a set of sentences "redundant" when a proper subset of this set describes the same situation without any loss of information. The first step in dissolving the One Over Many argument consists in simply recognizing that we can and should reduce our description to a proper subset. This smaller set must be complete but minimal. A set of sentences is, relative to a given situation, complete and minimal when it describes the situation fully and without redundancies. Thus, it is "minimal" in the sense that no sentence can be eliminated without doing harm to the completeness of the description.[1] The following three sets of sentences are equivalent and minimal in this sense:

S1: {(5), (7)}
S2: {(2)}
S3: {(3)}

We could say that S1, S2 and S3 are equivalent theories in the informational sense. Each set contains sentences that, taken together, imply all

1. The notion used here of a minimal informational set can be seen as an inversion of the notion of a maximally consistent set in logic, replacing consistency with informationality: adding even a single sentence to the maximally consistent set makes it inconsistent. Similarly, for minimal informational sets: if we eliminate even a single sentence of the minimal set, its information becomes informationally incomplete.

the others. No one is more informative than the others. They are mutual paraphrases. Note that here I use a holistic notion of paraphrase: I understand a paraphrase as not just a relation between two single sentences (that is, as just a particular case), but rather as a relation between two sets of sentences. Interestingly, according to the quantificational commitment criterion, S3 *prima facie* commits us to the existence of universals (it contains the sentence "*a* and *b* have a common property *F*" that quantifies over F), while S2 and S1 do not. Now, how can we eliminate the commitment of S3? Or, why should we prefer S1 or S2 to S3? And before this, why should we not take the complete union theory {(2), (3), (5), (7)} or even the maximal theory with all sentences (1)–(7) with all its commitments? Once we see the redundancy of the complete set of all sentences, the first step consists in reducing the complete theory to a minimal version of it. In this way, we reduce the commitment of all sentences to a small subset of it. Note, e.g., that {(1), (5)} is not an adequate set. It leaves open the possibility that Fa, $\neg Fb$, Ga, Gb, i.e., a situation in which *a* and *b* have a common property that is not *F*. Thus, it is not complete from the informational point of view. It does not contain the information that *b* is *F*.

What is the justification of this first step? Any nominalist would probably say that in ontology less is more. We should avoid unnecessary commitments. But, in fact, I think that reduction is not exclusively a task of nominalists. Anyone who tries to decide the question of the basic categories of reality, nominalist and Platonist alike, searches for reductions. The Platonist who proposes the bundle theory of particulars, according to which particulars are nothing but a bundle of properties, is reducing one category to another. He is a Platonist because he reduces particulars to universal properties, and not the other way round. But, supposing that many different reductions are possible, which should we prefer? At this point, Goodman and Quine (1947) are very honest and claim that the nominalist preference for renouncing abstract entities is a "philosophical intuition that cannot be justified by appeal to anything more fundamental". I think, at this point, again, that we can go beyond Quine and propose a better answer.

Indeed, I think that elimination is not an exclusive principle of nominalism. In its essence, ontology is reductive.[2] But a second reason can be added. If we do not require non-redundancy, an exponential increase in

2. See arguments for this in Imaguire (2008).

ontological commitments becomes inevitable. Take (5) alone: *a* is *F*. If this is true, the following sentences are arguably true for the Platonist:

(8) the binary relation of instantiation holds between *a* and *F*
(9) the ternary second order relation of instantiation holds between *a*, *F* and the binary relation of instantiation
(10) the second order property of being instantiated by *a* is instantiated by *F*.

and so on *ad infinitum*. Now, applying the quantificational criterion (or the Fregean principle that the singular term of a true sentence must refer) we could apparently derive the conclusion that entities like *the binary relation of instantiation*, *the ternary second order relation of instantiation*, *the second order property of being instantiated by a*, and infinitely many more, really do exist.

I propose that we should decide which sets are minimal sets in some intuitive "informational" sense. This can be made more precise by the new ostrich. The "new ostrich" is an ostrich who goes beyond Quine, the respectable "old ostrich". Basically, two approaches seem to be possible. The first works in terms of the entailment of sentences. We could take the complete set of sentences and investigate which sentences contain other sentences. This containment does not have to be logical or analytical. We could assume something like inferential semantics and accept that (1)–(7) can be logically (in a broad sense) inferred from each of the minimal sets.

The second approach does not appeal to language. It is more ontological and uses the notion of the truthmaker, in particular a rule that could be called a principle of "truthmaker minimalism". Take the two facts that *a is F* and that *b is F*. These two facts together are sufficient to make (1)–(10) all true. In typical truthmaker terminology: the facts *Fa* and *Fb* necessitate the truth of (1)–(10). They necessitate, further, all truth-functional derivable sentences like, "$\text{Fa} \vee \text{Fb}$" or "$\text{Fa} \rightarrow \text{Fb}$". Indeed, I think that (2) is just such a degenerate case. In this way, we can even accept that (8) is true. But then, one single truthmaker suffices to make infinitely many sentences true. This would not only make the ontological commitment too generous, but would even make the decision impossible from a practical point of view. How can we decide the ontological commitment of infinitely many sentences? If all the infinitely many sentences were construed in a simple recursive manner, one could even imagine a decidable set of their ontological commitments. But language is too misleading, and as soon as we accept any kind of re-description, the ontological commitment becomes unmanageable.

Let us consider a different example, EX2, in order to illustrate step 1. This example makes it clear that truthmaker analysis is fundamental for establishing minimal sets. It will also be important for explaining our next step. Take the following sentences as true:

(1*) Peter is 20 years old.
(2*) John is 30 years old.
(3*) John is older than Peter
(4*) Peter is younger than John
(5*) The age difference between Peter and John is 10 years
(6*) John is 10 years older than Peter

Sentence (5*) expresses something about *the age difference between Peter and John*. By existential generalization and applying the quantificational criterion, we should conclude from (5*) that there is an entity *the age difference between Peter and John*. It is obvious that by merely re-describing the same situation we can introduce infinitely many new entities into our ontology. If we clean up the redundancies, however, we obtain just a few minimal sets:

C*2: {(1*), (6*)}	C*3:{(2*), (6*)}	C*4: {(1*), (3*), (5*)}
C*5: {(2*), (3*), (5*)}	C*6: {(1*), (4*), (5*)}	C*7: {(2*), (4*), (5*)}

and my favorite

C*1: {(1*), (2*)}

According to truthmaker minimalism: the facts that Peter is 20 years old and John is 30 years old are all the truthmakers we need. Taken together, they necessitate the truth of (1*)–(6*) (and of many other derivative truths). Each set of complete but minimal sets of sentences establishes enough truthmakers to make all the redundant sentences true. Of course, Quine would avoid arguing in terms of facts necessitating truths—but, as I adverted, this is "new", not "old" ostrich nominalism.

Our final ontology will only be fully determined after the next two steps. Nevertheless, already in the first step we gain substantiality. The requirement of minimalism, on the one hand, avoids unsubstantial Platonism, i.e., multiplying entities without reservation. The requirement of completeness of description, on the other hand, avoids unsubstantial nominalism, i.e., eliminating entities that should not be eliminated.

2.2 *Second step: grounding*

The first step explained why we should not extract ontological commitments from redundant theories like {(2), (3), (5), (7)}. But given that there are various different minimal sets like S1, S2 and S3, the work is not over. Which of them should we select to submit to the commitment test? Let us remember that according to Quine's own commitment criterion, when translated into second order logic, S3 commits us to the existence of universals. Why should we not simply select S3 and conclude that universals exist? Why should we prefer S1 or S2 to S3?

At this point, Quine seems to be somewhat evasive: "One may admit that there are red houses, roses, and sunsets, but deny, except as a popular and misleading manner of speaking, that they have anything in common" (1948, 81). Why does the sentence "there are red houses, red roses, red sunsets" have the status of a "prephilosophical common sense in which we must all agree", while "the houses, roses and sunsets have something in common" is just a "popular and misleading manner of speaking"?

The best Quine offers here is an epistemological, not an ontological answer. He objects that the introduction of universals like redness represents no gain in explanatory power (1948, 81).

> That the houses and roses and sunsets are all of them red may be taken as ultimate and irreducible, and it may be held that McX is no better off, in point of real explanatory power, for all the occult entities which he posits under such names as "redness". (1948, 81)

Quine's general attitude in his ontology is clear. If we have different sentences that can be considered mutual paraphrases (such as "*a* and *b* are *F*s" and "*a* and *b* have *F*-ness in common"), one that commits us to the existence of X and one that does not, we are free to prefer the one requiring less ontological commitment. Or, perhaps better: we are allowed to select the sentence with the commitments that best match our ontological preferences. Indeed, I think this is correct: if the nominalist can show that we do not need to assume the existence of universals, the task is already accomplished. We do not have to show that the realist assumption of existence is false, but only that it is not necessary.

But the new ostrich nominalist can go beyond Quine and make a point to fill this gap. There is a good reason why S1 is really objectively better than S2 and S3. A new ontological device that was not available to Quine has been introduced by contemporary metaphysicians and is useful here, viz. the notion of ontological *grounding*. Actually, there is a family of correlated notions: *grounding*, *primitiveness* and *ontological dependence*. These three notions are, of course, different, but closely connected. In principle, these notions are intuitive and have been used in an intuitive way. Much work is still needed to make them sufficiently precise. In any case, it is true that we have some intuitions concerning primitiveness and fundamentality. And these metaphysicians appeal to this intuition: that Socrates is more fundamental than the single set with Socrates (Fine 1994), that a particular is arguably more fundamental than its constitutive tropes (Mulligan, Simon & Smith 1984), that an event is more fundamental than its temporal parts (Mulligan & Smith 1986), etc. I strongly believe that Quine would not be sympathetic to this ontological notion—but the new ostrich is free not to slavishly follow the old ostrich each step of the way.

Let us begin by illustrating this step with our EX2. Both the sets C*1, with the members

(1*) Peter is 20 years old.
(2*) John is 30 years old.

and C*6, with the members

(1*) Peter is 20 years old.
(6*) John is 10 years older than Peter.

are minimal in my sense. If Peter is 20 years old, and John is 30 years old, then John is 10 years older than Peter. According to the principle of truthmaker minimalism, we do not need truthmakers for (1*), (2*) and (6*). Truthmakers for (1*) and (2*) are enough truthmakers. If (1*) and (2*), then (6*) is automatically true. But, similarly, if Peter is 20 years old and John is 10 years older than Peter, it follows that John is 30 years old. Indeed, from a semantic or logical point of view, no one set has priority over the others. But this is exactly the heart of the ontological intuition of grounding: it is clear from the ontological point of view that John is 10 years older than Peter *in virtue of* Peter being 20 years old and John being 30 years old, and not the other way round. Exactly the same holds for the facts expressed by (3*), (4*) and (5*).

The fundamentality of (1*) and (2*) also becomes clear when we compare all minimal sets. In every set, one of them is elementary. More: it would be strange to suppose that (1*) is primitive and (2*) is not, for both have the same logical form and express exactly the same kind of fact. And some principle of equivalence must be supposed: facts of the same form must have the same ontological status. Thus, all minimal sets but C*1 are derivative.

Let us come back to the One Over Many argument and the sets S1, S2 and S3. That S1 is preferable to S2 is not really relevant, since S2 also does not commit us to the existence of universals. But it certainly seems reasonable to suppose that S1 is more fundamental than S2. The conjunction *Fa* & *Fb* is true, *because Fa* is true and *Fb* is true, and not the other way round. The conjunction is true *in virtue of* the conjuncts being true. The atomic facts are the *grounding reason* for the conjunctive fact.

What about S3, the set with the sentence

(3) *a* and *b* have a common property *F*?

Quine would deny that there is such a second order fact. The new ostrich has an argument to support this denial: it is the case that *a* and *b* have a common property *F,* because *a* is *F* and because *b* is *F*, not the other way round. At least in the case where F is an intrinsic property, *a* can be *F* independently of *b* being *F* or even of *b* existing at all. The particular *a* could be *F* without *b* being *F*. The same holds for *b*: it can be F independently of the nature and existence of *a*. Property agreement is a derivative fact, not a grounding one.

One could argue that there are some exceptions to this analysis. Suppose that *a* and *b* are human beings, and *b* is the son of *a*. It is arguable that *b* is a human being just *because* he is an offspring of *a*, and *a* is a human being. Thus, the exemplification of humanity by *b* is grounded on (or "derived from") the exemplification of humanity by *a*. Because of cases like this, I use the limiting expression "at least in the cases of intrinsic properties". But even in this case, I would reply that the sense of "because" in which *b* is *F because a* is *F* is a biological sense, not an ontological grounding one. From an ontological point of view, *a* and *b* are *F* in virtue of nothing more.

2.3 *Third step: ontological import*

Step 1 gave us various minimal sets of true sentences, and step 2 helped us to select just one of them, viz. the most fundamental one. Finally, in this third and final step we derive the ontological commitment of the theory by analyzing the ontological import of each of the sentences of this set.

Take EX1 and its fundamental minimal set C3. It has two members:

(5) a is F
(7) b is F

What is the ontological import of (5) and (7)? Since they are absolutely similar, I conclude that the analysis of one of them will be immediately applicable to the other. Thus, let us take "a is F". Well, the existence of *a* cannot be denied. Indeed, Quine's proposal of ontological commitment should be understood as a principle of theoretical coherence. We cannot say, in the same breath, *Fa,* and *a* does not exist. For, if "Fa" is true, then *there must be* an *x* such that *x* is *F*.

But, given that "Fa" is true, do we have to conclude that *F*-ness exists? After all, this is what is really at stake here. The answer of the ostrich is clearly "we do not have to". The best way to justify the negative answer presents a dilemma. If "Fa" commits us to *F*-ness, then either (i) "Fa" commits us to *F*-ness by force of a paraphrase, or (ii) "Fa" commits us to *F*-ness by its own force. The first horn is by paraphrase: If *F*-ness is an ontological import of "Fa" through a paraphrase, the question arises: through which paraphrase? The most natural and plausible answer is, of course, something like

(11) F-ness is instantiated by a.

But why is this alternative wrong? Simply because arguing for the existence of *F*-ness based on (11) is to surreptitiously add a new sentence to the minimal set. The whole procedure consists precisely in selecting, at the first step, a minimal set of sentences. In EX1 we had {(5), (7)} and not {(5), (7), (11)}. And that we should accept only minimal sets is already argued above. In particular, by accepting non-minimal sets we open the door to an infinite multiplication of entities: not only must *F*-ness be accepted, but infinitely many other entities. The only non-arbitrary options are either: only *a* exists, or infinitely many other entities exist (and this just because *a* is *F*!).

The other horn of the dilemma is: *F*-ness is an ontological import of "Fa" by its own force. But this cannot be accepted in virtue of the principle of substantiality. Only unsubstantial Platonism would defend the view that the mere occurrence of the predicate "F" in a true sentence "Fa" is sufficient to commit us to the existence of the universal of *F*-ness. All sentences contain predicates and, thus, we would include infinitely many properties in our ontology. And even most realists, like Russell and Armstrong, are unwilling to accept this. As we saw above, to exercise "ontological correctness", even Platonists add some additional requirement for predicates to obey the principle of substantiality.

According to Quine, predicates are syncategorematic expressions and not names of properties. Thus, in order to explain the truth-conditions of "Fa", we do not need to assume that "F" refers to something. Platonists reply to this with the One Over Many argument: if it is true that "Fa" and that "Fb", then we must concede that "a and b have F-ness in common" is also true. And in order to explain the truth-conditions of this sentence, we need to assume that "F-ness" refers to something—a universal. But, as we saw in these three steps, the ostrich nominalist can, like Quine, simple say that "a and b have something in common" is a "popular and misleading manner of speaking" (1948, 10).

To sum up: the ontological import of "Fa" is only *a*, the import of "Fb" is only *b*. Thus, the commitment of the minimal set C3 and the commitment of the original set (1)–(7), are the particulars *a* and *b*, and nothing more. From this fact alone, viz. the fact of property agreement, we should not conclude the existence of universals. By the way, due to parity of reasoning, the result for EX2 is the same. We selected C*1 as the most adequate minimal set. It entails the sentences "Peter is 20 years old" and "John is 30 years old". Thus, the final ontology of (1*)–(6*) is composed of Peter and John, but not of the properties *to be 20 years old*, or *to be 30 years old*.

Conclusion

We began with two problems for the ostrich: the first was that the ostrich does not take predicates with ontological seriousness; the second was the paraphrase problem. The solutions are, I believe, now clear.

Is it true that the ostrich does not take predicates with ontological seriousness? We can now definitely conclude that it is not true. As we have seen, there are two different ways of giving a substantial answer to this question: substantial Platonism and substantial Nominalism. And to assure substantiality, even the Platonist must accept some additional principle for deciding whether a given predicate commits us to the corresponding universal. For some, only primitive irreducible and non-logical predicates commit us to universals. For others, only predicates of basic laws of nature commit us to universals. The ostrich also has an additional criterion: only predicates we quantify over commit us to universals. The ostrich solution is, as we saw, among the most popular alternatives, the most substantial one. It is evident that every theory has primitive predicates, and that we need predicates for formulating scientific laws. But that we have to quantify over predicates in the minimal set of more fundamental true sentences is certainly a more substantial criterion.

What about the paraphrase commitment problem? Given sentences S1 and S2, since both are synonymous, why should we select the commitment of one of them and not that of the other? Indeed, Quine never answered this question, and we have undertaken to provide an answer here. One possible approach would be, of course, to deny that both are synonymous (see e.g. Simons (1992, 152). Indeed, Quine himself warned us to be careful with the notion of synonymy. In the case of the sentences (1) to (7), we have some alternative sets of sentences that are equivalent in their information. Quine's suggestion seems to be: just take the sentence with the least ontological input. The new ostrich's suggestion is more substantial: just take the sentences (or the set of sentences) that express *grounding* or *more fundamental* facts. Thus, as everyone suspected, the ostrich really does hide his head in the sand—not because he is afraid of a metaphysical problem, but simply because he is looking for grounding relations in ontology. And nothing is more fundamental than predication. Note that this solution is—quite in a Quinean spirit—holistic. We should not take sentences in isolation, but rather find sets of sentences that constitute a whole theory. And given such a set, it will never be the case that two synonymous sentences would be elements of it—at least if we restrict ourselves to a minimal set, as the new ostrich proposes.

This solution also implies a direct answer to question F3: "In virtue of what is the particular *a* an *F*?" The straightforward answer of the new ostrich is: in virtue of nothing more fundamental: *a* is *F*, full stop. I think this is also the answer of the "old" ostrich. Quine says, "that the houses and roses and sunsets are all of them red may be taken as ultimate and irreducible" (1948, 81). What could "ultimate and irreducible" possibly mean other than "in virtue of nothing more fundamental"? Indeed, some have proposed to reduce predication (actually better: instantiation) to other basic relations: participation, (genuine linguistic) predication, set theoretical membership-relation or similarity. Of course, as Lewis (1983) puts it, each theory has the right to suppose its own primitives. The ostrich simply takes facts like *Fa* as primitive. And there are many good reasons for this.

What about Armstrong's favorite formulation (F2): how can different particulars be identical in nature? Indeed, it is interesting to note that for the resemblance nominalist, the simultaneous instantiation of the same property by different particulars is simply the most fundamental fact. It has no grounding reason. On the contrary, it is the grounding reason for everything else, even for predication. In principle, such an answer is as legitimate as any other. For, as we mentioned in the last paragraph, every theory has the right to suppose its own primitives. But for the ostrich, the fact that different particulars can instantiate the same property is not fundamental. It is derivative: it is not necessary for a particular *a*, in order to be *F*, to be similar to anything else (to another instance of *F*). And so, if two particulars are both *F*s, it just happens that different particulars are "identical in nature". For the ostrich, the predicative nature of properties is enough to explain this. Again, the ostrich puts his head in the sand not because he is afraid of metaphysics, but because he keeps his eyes open for grounding relations, i.e., for what is really fundamental.

References

Alston, William 1958: "Ontological Commitments". *Philosophical Studies* 9, 8–17.

Armstrong, David 1978: *Universals and Scientific Realism. Volume I: Nominalism and Realism.* Cambridge: Cambridge University Press.

— 1980: "Against 'Ostrich' Nominalism: A Reply to Michael Devitt". *Pacific Philosophical Quarterly* 61, 440–449.

Campbell, Keith 1990: *Abstract Particulars*. Oxford: Basil Blackwell.

Devitt, Michael 1980: "'Ostrich Nominalism' or 'Mirage Realism?'". *Pacific Philosophical Quarterly* 61, 433–439.

Donagan, Alan 1963: “Universals and Metaphysical Realism”. *Monist* 47(2), 211–246.
Fine, Kit 1994: “Essence and Modality”. *Philosophical Perspective* 8, 1–16.
Goodman, Nelson & Quine, Willard van Orman 1947: “Steps Toward a Constructive Nominalism”. *Journal of Symbolic Logic* 12, 105–122.
Hirsch, Eli 2009: “Ontology and Alternative Languages”. In: David Chalmers, David Manley, & Ryan Wasserman (eds.), *Metametaphysics*. Oxford: Oxford University Press, 231–259.
Imaguire, Guido 2008: “Ockham’s Razor and Chateaubriand’s Goatee”. *Manuscrito* 31(1), 139–154.
Lewis, David 1983: “New Work for a Theory of Universals”. *Australasian Journal of Philosophy* 61(4), 343–377.
MacBride, Fraser 2005: “The Particular-Universal Distinction: A Dogma of Metaphysics?”. *Mind* 114, 565–614.
Melia, Joseph 2005: “Truthmaking Without Truthmakers”. In Helen Beebee & Julian Dood (eds.), *Truthmakers: The Contemporary Debate*. Oxford: Clarendon Press.
Mellor, David & Oliver, Alex 1997: *Properties*. Oxford: Oxford University Press.
Mulligan, Kevin & Barry Smith 1986: “A Relational Theory of the Act”. *Topoi* 5(2), 115–130.
Mulligan, Kevin; Simons, Peter & Smith, Barry 1984: “Truth-Makers”. *Philosophy and Phenomenological Research* 44, 278–321.
Oliver, Alex 1996: “The Metaphysics of Properties”. *Mind* 105, 1–80.
Peacock, Howard 2009: “What’s Wrong with Ostrich Nominalism?” *Philosophical Papers* 38 (2), 183–217.
Quine, Willard van Orman 1948: “On What There Is”. *Review of Metaphysics* 2. Reprinted in Willard van Orman Quine, *From a Logical Point of View*. Cambridge, MA: Harvard University Press.
Ramsey, Frank 1925: “Universals”. *Mind* 34, 401–417.
Rodriguez-Pereyra, Gonzalo 2000: “What Is the Problem of Universals”. *Mind* 109, 255–273.
Simons, Peter 1992: “Ramsey, Particulars and Universals”. *Theoria* 57, 159–161.
Van Cleve, James 1994: “Predication Without Universals? A Fling with Ostrich Nominalism”. *Philosophy and Phenomenological Research* 54 (3), 577–590.

Grazer Philosophische Studien
89 (2014), 205–228.

QUINEAN WORLDS: POSSIBILIST ONTOLOGY IN AN EXTENSIONALIST FRAMEWORK

Pedro SANTOS
Universidade Federal de São Paulo

Summary

In "Propositional Objects" Quine sketched a construction of possible worlds that is consistent with his extensionalism. If this construction succeeded, the extensionalist would be able to make good sense of intensional notions and objects. However, Quine ended up rejecting the proposal due to the problem of "trans-world identity". In this paper I start by reviewing Quine's construction and suggesting some modifications. I then defend it against Quine's objection. I also defend it against two more general objections raised by David Lewis. I conclude that this is the modal metaphysics an extensionalist such as Quine should adopt.

1. *Introduction*

Quine was both an *extensionalist* and a *modal eliminativist*. That is to say, he thought both that the best canonical language for total science is an extensional language, and that there is no sufficiently clear notion of necessity. It might seem at first as if modal eliminativism follows immediately from extensionalism. For aren't modal notions paradigmatic non-extensional notions? However, this first impression is misleading. Elimination is not the only way to deal with a notion you find problematic. You might also attempt to *reduce* it to more acceptable ones. In particular, extensionalists might attempt to "extensionalize" (purportedly) non-extensional notions.

Quotation provides a good illustration of the extensionalization strategy. Quotation appears to create non-extensional contexts: from "Cicero' has six letters' and 'Cicero is Tully', it does not follow that "Tully' has six letters'. However, once quotation is analyzed in terms of spelling, as Quine (1960) following Tarski suggests, these failures of substitutivity become merely apparent or a mere "surface phenomenon" easily "regimented away".[1] In this way quotation is not eliminated, but shown to be extensionalistically acceptable. Now, one might attempt to apply the same strategy to modal notions. For example, if we analyze necessity as truth in all possible worlds, then *provided* the notion of a possible world can be made acceptable to the extensionalist, we will end up with an extensionalistically acceptable notion of necessity.

In a sense this is what the so-called *modal realism* of David Lewis tries to accomplish. Lewis' view is, arguably at least, both extensional and reductive (hence not eliminative) about modality. But, of course, Quine did not accept a plurality of worlds in Lewis' sense. And so one might suggest that eliminativism follows from extensionalism, *together with* Quine's ontological scheme of *this-worldly* concrete particulars and classes. Modal realism would then be the only way to reconcile extensionalism and modality. But actually that is not so clear either.

In "Propositional Objects" (1968) Quine himself considered an account of possible worlds that is both extensional and consistent with his ontological views. Very roughly, on this view worlds are identified with certain sets that are taken to describe or represent ways the world could be. As is well known, once we have a reasonable theory of possible worlds in place we can define not only modal operators but also useful notions of property, relation and proposition. However, in spite of this huge potential pay off, Quine ended up rejecting the proposal rather quickly, for reasons related to trans-world identity. Later this kind of view was further developed and criticized by Lewis (1986) under the name of *linguistic ersatzism*. Lewis raised two problems that have become standard: the problem of reduction and the problem of expressive power.

In the present paper, I shall argue that something close to Quine's original idea can be defended against these objections (both Quine's and Lewis'). I call the resulting view *Quinean linguisticism*. My point will not be that this is the *best* metaphysics of modality available. It will be the rather more modest one that, though not without its difficulties, Quinean linguisticism provides an account of the apparatus of *possibilia* that is reasonable and consistent with Quine's ontology and ideology. (I also think it is preferable to modal realism; but I will not argue this point here.) If this is true, then it is hard to resist the conclusion that, in embracing modal eliminativism, Quine did not give his brand

1. See Quine (1960), 143. For the same idea in Tarski see, for instance, Tarski (1956).

of extensionalism its due, making it look much less palatable than it has to be.[2]

2. *Worlds in extension*

In (1968) we find Quine sketching a construction of possible worlds that, as he puts it, "stays within a clear extensional ontology" (152). The notion of an extensional entity may not be a very clear one; but luckily for present purposes we don't need to go into its analysis. What is essential is that no appeal will be made to typical intensional entities such as properties and propositions. Of course, the construction is also supposed to stay within an extensional *ideology*, in the sense that ultimately no non-extensional notions such as modal operators are employed.

Quine's immediate goal was to construct "states of affairs" as sets of possible worlds and then take them to be the objects of at least some "primitive cases" of the propositional attitudes, such as a cat's desire to get on a certain roof. What the cat desires is the set of all worlds in which he is on the roof. But it is clear that if his extensional worlds can serve this particular purpose, there are other significant purposes they can be made to serve as well. First, belief in general could, to some extent at least, be treated in this way. Quine in fact acknowledges this point, if only implicitly, by applying the apparatus to ordinary human beliefs as well. More importantly for us, necessity could be defined as truth in all possible worlds; n-adic relations could be defined as functions from possible worlds to sets of n-place sequences of possible individuals; and in general all the analyses afforded by possible worlds would seem to be available to an extensionalist.

Now, as remarked above, Lewis' theory of worlds can also be said to be acceptable to the extensionalist as such. For worlds, conceived as concrete universes, would seem to be extensional entities. Moreover, there are no non-extensional operators in Lewis' basic conceptual repertoire. But, of course, Quine did not believe in a plurality of concrete worlds. For Quine, the actual world is the only world there is; and it consists of physical objects

2. I shall not discuss Quine's classic objections to modal notions. If they showed such notions to be incoherent, then the present exercise would lack point. But modal notions are by now firmly established as legitimate notions and Quine's objections are widely considered to fail. See in particular Smullyan (1948). This only makes it all the more urgent to know whether or to what extent Quine's actualist extensionalism can accommodate modal notions.

(in a broad sense, comprising the "material contents" of any space-time region) and classes. How does he propose to accommodate possible worlds within such an austere ontology?

Quine's idea is a version of what Lewis dubbed *ersatzism*. The ersatzist's strategy is roughly as follows. He starts with an intuitive domain of all possible worlds. He then associates each world with a different "complete description" of it. Finally, he drops the worlds intuitively conceived and thinks of the descriptions themselves as being the worlds. Propositions are often taken to play the role of complete descriptions. A *world proposition* is one such that, possibly, it is true and implies all truths.[3] Quine's problem is to find, within the limited means available to the extensionalist, things capable of playing the role of complete descriptions.

To see how the extensionalist might attempt to solve this problem, let's consider a simple scenario introduced by Quine. Let's pretend for a moment to hold a Democritean physics in which everything is made of "atoms" which are "homogeneous in substance and differ only in size, shape, position, and motion". Let's also pretend to hold a Democritean *metaphysics*, according to which, necessarily, everything supervenes on the properties of the atoms. Then, once we have settled, for each point of spacetime, whether or not it lies inside an atom, we have settled everything. A possible world can then be defined as an assignment of "occupied" or "empty", but not both, to each spacetime point. Equivalently, a possible world can be seen simply as a set (or, if you prefer, aggregate) of space-time points, intuitively those space-time points that are occupied according to the world.

Since possible worlds in this sense are actually existing "abstract" objects, and since they represent by convention or stipulation, Lewis (1986) classifies this kind of view as a form of *linguistic ersatzism*. Here it is important to notice that we are thinking of languages in a generous sense. The expressions of the language need not be linguistic symbols in any conventional sense. In particular anything can be its own name. This can be seen to constitute an advantage over Carnap's (1956) forerunner notion of state-description, in which language appears to be conceived more narrowly.

This account is also sometimes considered a form of *combinatorialism*.[4] However, in spite of some evident similarities, this classification doesn't

3. See (Prior & Fine 1977).

4. Armstrong, for instance, says that in "Propositional Objects" Quine "toyed" with combinatorialism (Armstrong 1989). Besides Armstrong's own, other versions of combinatorialism include (Skyrms 1981) and (Bigelow 1988). The idea is that the range of possibilities is the range of "recombinations" of actually given things and traces back, of course, to Wittgenstein's *Tractatus*.

seem entirely appropriate. First, the idea that all combinations represent possibilities is not an essential part of the view. Indeed, if we think of "occupied" and "empty" as the two fundamental states of spacetime points, taken as fundamental individuals, then presumably we will wish to rule out assignments that attribute both states to the same individual. In the present case this problem need not arise, since we can define "empty" as "not occupied" and let the absence of a point in a world represent its emptiness; but of course this is due to the particularities of the simple Democritean scenario.[5] Also, we are not thinking about these "states" as being entities available for recombination in any literal sense. The extensionalist does not believe in properties as basic entities.

Notice that although Democritean metaphysics is a modal thesis, the definition of possible world itself is non-modal. We want the descriptions to correspond one-one with the worlds, intuitively conceived; and in setting up this correspondence we are free to make use of our beliefs about what is possible. We must only take care that no modal concepts appear, explicitly or implicitly, in the definitions we give. Of course, the fewer controversial assumptions are required to show the correctness of a definition, the better. It must be admitted that the extensionalist will not be able to score very high on this count; but I shall suggest below that the situation is not as bad as the simple example above suggests.

Of course, the Democritean physics/metaphysics is merely illustrative. The hope is that by modifying and complicating the descriptions something more realistic can be obtained. However, already at this illustrative level several questions arise. For example, imagine a world consisting of just two perfectly similar atoms, A and B, at rest. It seems possible for them to trade places. So we seem to have two worlds here. However, since each of these worlds has exactly the same occupied points, we have only one assignment corresponding to both. Thus even assuming the Democritean framework there is still a question whether all possibilities are being covered. There are also doubts about whether this kind of theory can really provide a reductive account of necessity. For even though possible worlds were defined in a purely extensional way as sets of points, in order to

5. It must be granted, however, that many, perhaps most self-proclaimed combinatorialists do introduce restrictions to the range of recombinations. Skyrms (1981), for instance, says that not all recombinations "count as possible" but does not attempt to explain how recombination is restricted. And Bigelow (1988) says that the "only" way to develop combinatorialism is "to introduce constraints on permissible combinations". Notice that they have in mind restrictions more substantial than those imposed simply by the ontological category of the elements.

define necessity we need to speak of something being the case *in* or, more appropriately in the present case, *according to* a world; and it is not entirely clear whether this locution can be explained in extensional terms. What is it, say for a talking donkey to exist according to a world, if not for that world to be such that if it accurately described the way things are, i.e. if it *were* actualized, then a talking donkey *would* exist?

We shall come back to these questions later on. In the remainder of this section I would like to discuss some of the complications Quine introduces and then consider how to arrive at a more realistic theory.

In (1968) Quine is worried, in the first place, about the reification of space-time. He didn't want to have "two sorts of individuals", namely space-time points and "portions of matter" (1968, 148) in his ontology. So he proposes to "by-pass" the points by adopting a system of coordinates and speaking of quadruples of real numbers. Instead of a class of space-time points we would see a possible world as a class of quadruples of real numbers. But it is not entirely clear that his move will work. How can we fix a coordinate system other than by specifying a *point* of origin? And if the quadruples of points represent spacetime points, aren't we committed to their existence anyway, even though they are no longer constituents of the descriptions? If we wish to avoid the two sorts of individuals, perhaps it is preferable to choose spacetime points.[6]

One further advantage Quine sees in thus "by-passing" space-time points is that we thereby free ourselves to adopt a relativistic view of space-time. But this is controversial. Quine seems to be assuming that a substantivalist view about space-time implies a Newtonian "absolute" space and "absolute" time. But according to Sklar (1977) these issues are largely independent. In particular, the adoption of relativistic physics does not imply the adoption of a relationist view of space-time. Discussing the Minkowski spacetime appropriate for special relativity, he says: "Spacetime has its independent existence and structure, as space and time did before. So, for example, a spacetime totally empty of all events is perfectly intelligible in this view." (Sklar 1977, 163) Later on he says that in the case of general relativity, the picture is more complicated, but "not in such as way as to make the substantivalist account prima facie untenable" (Sklar 1977, 164). Unfortunately, this topic is too large and complex for us to pursue it any further here.

6. Quine considers this possibility in (1976a) and in (1981). 'Portion of matter' is a peculiar notion to take as basic anyway. The usual alternative is to work with point-particles, or time-slices thereof.

There is also a complication that arises from Quine's decision to work with coordinate systems. There are many ways in which a coordinate system can be fixed; and clearly two classes of quadruples, each with its own coordinate system, can be so related that they should be seen as representing the same possible world. For the quadruples may compensate for the difference in the origin of the coordinates. Here one might see another reason for speaking directly of points as basic individuals. But Quine chooses to complicate his theory and equate a possible world with a certain equivalence class of classes of number quadruples. He ends up, roughly, with a possible world being a class containing all the classes "geometrically similar" to some class of number quadruples, where two such classes are geometrically similar if one can be converted into the other by changing the axis and multiplying all numbers by a constant factor.

From now on I shall avoid these complications by assuming a domain of basic individuals of unspecified nature. I assume only that the individuals in this domain are atomic, i.e. without proper parts, and momentary. So they could be spacetime points or momentary stages of point-particles. This is controversial, of course; but presumably what follows could be adapted to a view in which there are no atomic individuals and no temporal parts.

Now here is a suggestion of how one could arrive at a more realistic version of this view. It is not implausible to assume that there is some domain of simple properties that actually play the role played by "occupied" and "empty" in the Democritean example. That is to say, properties such that, necessarily, everything supervenes upon their distribution over individuals. (Keep in mind that, in spite of the Democritean example from which we started, we are not assuming that this base should be "materialistic" in any sense, although some people might want to make that assumption.) If the extensionalist believed in properties, he could take these properties themselves, whatever they are, as constituents of the worlds; but since he doesn't believe in properties, he must work with predicates that intuitively stand for those properties. The identification of these predicates is an *a posteriori* task for "total science" and at any given point the extensionalist will work with the best available theory. Then a possible world could be defined as a set of pairs such that the first member of each pair is an m-tuple of basic individuals and the second member is an m-place predicate. The presence of any such pair in the set means that the predicate applies to the individuals; if a pair is absent, this means that the predicate does not apply to the individuals.

One might envision certain modifications to make this picture more plausible. For example, any felt metaphysical incompatibilities between predicates can be incorporated into the definition of possible world. We might impose conditions like: for any individual i, if a world has <i, F> as a member it does not have <i, G>. Something like this was already present in the definition of a world as an assignment of "occupied" or "empty", *but not both*, to spacetime points. So there is no need to assume the "modal independence" of the "atomic facts", although some people might want to make that assumption. Presumably, since we are dealing with fundamental states only, these restrictions will stay manageable. Also, we may want to allow pairs <c, F>, where c is a *complex* individual (a mereological sum of basic individuals) thus allowing for the possibility of "emergent" fundamental states that apply to complex individuals but cannot be reduced to a pattern of instantiation of fundamental states by fundamental individuals.

This construction rests ultimately on the choice of a theory of the fundamental empirical nature of the world. This reliance on an empirical theory might seem odd; but actually it is quite in keeping with Quine's naturalism. For Quine there is no privileged epistemological point of view external to the empirical point of view; all our beliefs, including our metaphysical beliefs, rest ultimately on experience. So it should come as no surprise that our theory of modality is based on our best theory about the fundamental nature of the empirical world.

From now on I shall call this view of how an extensional theory of possible worlds should be constructed 'Quinean linguisticism'. I think Quinean linguisticism is rather more defensible than it seems at first; and I will attempt such a defense below. For now I observe only that it does not amount to an identification of metaphysical necessity and physical necessity. For, taking Democritean physics as an example, there will be worlds in which the atoms behave in such a way that it will be natural to say they are being governed by different laws. But it is not physically possible for the physical laws to change. So these worlds are not physically possible. Still, this kind of view may be thought to excessively restrict the range of possibilities, thus "falsifying the facts of modality". We shall come back to this issue below.

3. *Quine's objection*

As already mentioned, Quine's original motivation for the introduction of possible worlds was to construct "states of affairs" as classes of possible worlds and to take them as the objects of propositional attitudes. For instance, if I believe that the Great Pyramid has equilateral faces, what I believe is the class of all possible worlds in which it does have equilateral faces. Quine sees this as settling "the matter of individuation" that afflicts the notion of proposition. It is unclear whether to count the *proposition* that the Great Pyramid has equilateral faces as being the same as the *proposition* that the Great Pyramid has equiangular faces, since propositions are individuated in terms of the obscure notion of *meaning*. But the worlds in which the pyramid has equiangular faces are, uncontroversially, the worlds in which is has equilateral faces. So the question is settled for states of affairs understood as sets of worlds. Quine ended up rejecting this idea and in this section we shall discuss his reason for doing so.

However, notice first that it is not at all clear what it means to say that w is a world *in which* the Great Pyramid has equilateral faces, or a world *in which* donkeys talk. The reason is that 'equilateral faces' and 'talking donkey' are presumably not going to be predicates of any fundamental empirical theory of the world, and so do not appear explicitly in worlds. Moreover, we cannot simply say that a donkey talks according to w if, had the distribution of fundamental states been as w says it is, a talking donkey would have existed. This may well be true; but it is a modal statement, and so if we say this we cannot claim the analysis to be reductive. This is the problem of "implicit representation" raised by Lewis and it is part of what I called above the problem of reduction.

A world must represent both *de dicto*, as when it represents the world as being such that there are talking donkeys, and *de re*, as when it represents the world as being such that the Great Pyramid (*that* specific thing), has equilateral faces. Both forms of representation generate puzzles for the Quinean linguisticist; but Quine only raises the problem of *de re* representation. In this section I shall assume that the problem of *de dicto* representation has been solved, i.e., I shall pretend we understand what it means for a world to be a world in which, say donkeys talk or pigs fly. Then, what remains to explain in an example such as the pyramid's is what it means for a world to represent something as being identical to the Great Pyramid. Later we shall come back to *de dicto* representation.

Quine presents his objection in the following passage:

> How is Catiline to be identified in the various possible worlds? Must he have been named "Catiline" in each, in order to qualify? How much can his life differ from the real life of Catiline without his ceasing to be our Catiline and having to be seen as another man of that name? Or again, how much can the pyramid differ from the real one? It will have to differ a little in shape if my belief about it [that the pyramid's faces are equilateral] happens to be mistaken. Is it sufficient, for its identification in other worlds, that it have been built by Cheops? How much then can his life differ from the real life of Cheops without his ceasing to be our Cheops? (Quine 1968, 153)

Quine doesn't even *attempt* to answer these questions. He simply starts weakening his account so as to avoid the issue altogether. I take this to mean that he sees such questions as having no answers and is inviting the reader to agree. But notice that there are in fact two problems Quine may be trying to raise with these questions. One is the problem of what it *means* to speak of something as existing in several possible worlds. This is basically a call for a theory of possible individuals. The other problem is the problem of how, if at all, we can *tell* that we have a case of trans-world identity. This is an epistemological problem that one should expect a theory of possible individuals to settle. Following (Divers 2002) I call the first problem the problem of trans-world *identity*, and the second the problem of trans-world *identification*. In the first case the charge would be that no sense can be given to statements of trans-world identity; and in the second that no satisfactory epistemology can be developed for these statements and that *de re* modal discourse is too loose to be worthy of serious consideration.

The problem of trans-world identity is clearly more fundamental, so we must start with it. The object of my belief that Catiline is a Roman is supposed to be the set of worlds in which he is a Roman. If we have no idea what it is for a world to represent something as being Catiline, then evidently we have no idea what it is for a set of worlds to be the set of worlds in which he is a Roman.

Now, if this is the problem Quine is raising, it is hard to avoid the feeling that he gave up too easily. The natural attitude would be for him to try and work out a theory of possible *individuals* and of trans-world identity to go together with his theory of possible worlds. He didn't do it; but he gave no specific reason why this can't be done. So we might try to fill in this gap for him. Now once we consider this question, the idea that naturally comes to mind is to apply the very same strategy Quine used in connection with a similar worry concerning the cross-*temporal* identification of material objects.

In (1950) Quine, echoing Heracleitus, asks how a man can bathe twice in the same river, given that new waters are always flowing in upon him. If the waters are not the same, in which sense can we talk of the same river? His answer is that a river is a "process through time" composed of many momentary "stages". We bathe in the same river twice by bathing in two stages of the same river. Generalizing this approach, we arrive at Quine's broad notion of physical object: a physical object is the "material content" of any space-time region. One way to make this a bit more precise is to think of a physical object as the *mereological sum* of the "matter" contained in a given region.

Now when we talk about an ordinary material object, such as a desk, we clearly don't have a precisely delimited physical object in mind. We can't say exactly when the desk started to exist, or when it is going to end; and at any moment in time we cannot say exactly where its spatial boundaries lie. We can put this by saying that the desk has vague boundaries (both in time and space); but on the present view it is really more a case of our having imprecise minds: entities in the world are precise, it is we who do not know precisely which physical object we are talking about. Thus in a later paper Quine says:

> What the vagueness of boundaries amounts to is this: there are many almost identical physical objects [...]. Any one of these almost coextensive objects could serve as the desk, and no one the wiser [...]. Nevertheless, they all have their impeccable principle of individuation [...]. (Quine 1981, 101)

It is not my intention to defend this "four-dimensional" view of material objects and temporal identity, or the semantic explanations of vague boundaries (though I find both appealing). Though controversial, these views are clearly live options embraced by many authors. My point is that an analogous solution to the case of trans-world identity seems possible. Roughly, we can think of possible individuals as being extended, not just across spacetime, but also across logical space, i.e., across possible worlds. Then physical objects will have not just temporal stages, but "modal stages" as well. If I think of my desk as an actualized possible object, then the thing in front of me is merely its actual modal stage or, as we shall say, its actual *manifestation* (or rather the present temporal stage thereof).

Let's try to work this out more precisely. The universe, according to the Quinean, contains basic individuals, mereological sums and set-theoretical constructions thereof. We have seen, in outline at least, how to find entities within this universe to play the role of possible worlds. Now we want to do the same for possible individuals. I suggest we start by thinking of the collection of *things* as comprising the basic individuals and all sums thereof. Things in this sense are roughly Quine's physical objects. Next we define a *trans-world individual* as a function f taking us from each possible world either to a thing or else to the empty set. The possible individuals will be the things, plus the trans-world individuals. If f is a *trans-world individual*, we call f(w) the *manifestation* of f at w.[7]

I assume that things exist in all possible worlds and that sums are *rigid*, in the sense that they necessarily have the parts they have.[8] The identity of things across possible worlds is not problematic: they are always the same sum of basic individuals. So we can just say that a thing has F in w iff w represents it as having F. (It is still unclear what it is for a world to represent something as having a non-basic property, but that is the problem of implicit representation to be discussed later.) Now a *trans-world* individual has a property F at world w iff f(w) has F in w, i.e., if w represents f(w) as being F. If the manifestation of a trans-world individual in a world is empty, we say it does not exist in that world. A *merely possible* individual is one that does not exist in the actual world. A possible individual is *actualized* if it exists in the actual world.

Thus consider again an ordinary individual such as my desk. Quine suggested we think of it as a sum of momentary entities, i.e. as a thing; but now the suggestion is that we identify it with a trans-world individual, with a function from possible worlds to things. In this sense, to say my desk is brown is to say that its actual manifestation is brown according to the actual world, which is equivalent to its really being brown. To say it could have been red is to say that there is a world such that its manifestation at that world is red according to it.

We can't simply say that the desk could have been red iff *this* thing is red according to some world. The problem is not that there are no possibilities for *this* sum, but rather that we would end up getting the wrong possibilities for the desk. Suppose, for instance, we are taking spacetime

7. Something along these lines appears in (Cresswell 1972), from which I borrowed the term 'manifestation'.

8. We could let basic individuals exist contingently, in which case sums would also be contingent; but they would still be rigid.

points as basic individuals. Then the desk, *qua* individual, is a certain spacetime region. Now there is no world according to which a spacetime region is differently located; but we want to say, of course, that the desk could have been differently located.[9] We face a similar problem if we work with particles. For we want to say the desk could have had slightly different parts; but mereological sums seem to be rigid in the above sense. And, of course, we want to say the desk could have failed to exist. Here existence is not to be taken in the sense of being identical to something. The desk has a manifestation in every world; since every world represents its manifestation as being identical to itself, every world represents the desk as being identical to something. In the sense of being identical to something we have *necessitarianism*: necessarily, everything necessarily exists. The desk exists contingently in the sense that it has an empty manifestation at some worlds. Things, on the other hand, having no manifestations, are not contingent in this sense either.

This may seem strange. When I ask if my desk could have been red, surely I am talking about the physical thing in front of me, not about a function of which the thing in front of me is a value? But compare: I point to something and ask whether *this river* passes through a distant town. Strictly speaking I am pointing to a spacetime segment of the river that clearly does not pass by the distant town. This is a case of "deferred ostension". I am really talking about a more inclusive thing, "the whole river", and asking whether it has a stage that passes through the town. So if I point to a desk and ask if something is a possibility for it, it is not so strange for me to be understood as talking about a more inclusive thing of which the thing in front of me is merely one temporal stage of one possible manifestation. Still, in non-modal cases we *can* be taken to be speaking of things rather than of trans-world individuals.

As before, doubts about individuation should be seen as arising from the vagueness of our concepts. Just as we normally don't know exactly which physical object we are talking about, we don't normally know exactly which trans-world individual we are talking about. We don't know, for instance, how much the material constitution of the desk could differ from its actual one. We can put this by saying that things have vague *modal* boundaries; again what this means is that we are imprecise when we talk about them. There are many objects that could serve as the trans-world desk and no one the wiser. It might be said that these possible objects differ from each other

9. See (Cresswell 1972).

more widely than the physical objects that could be taken as the physical desk. But that doesn't mean we impose no constraints. For example, there is no world where the desk has a manifestation that is both red and not red according to it, no world were it is red all over and green all over, no world where it is a basic individual, no world where it is conscious. Notice also that we actually feel pretty confident about the questions raised by Quine. We all agree that Catiline could have failed to be named 'Catiline', that the pyramid of Cheops could have been built by somebody else, etc. Therefore, it would be inappropriate to interpret 'Catiline' as meaning something that fails to have a manifestation in a world that is not called 'Catiline' there, etc.

Interestingly, Quine did not fail to consider this kind of theory, even though it is not mentioned in (1968). In "Worlds away" (1976) he proposes to discuss the suggestion that "identifying an object from world to possible world is analogous […] to identifying an object from moment to moment in our world" (Quine 1976, 859). Curiously, Lewis (1986, 217) claims that this paper appears to be about "genuine modal realism". His reason is that Quine does not explicitly connect the discussion there with the extensional worlds of (1968). However, it seems more plausible to me that Quine was taking his earlier work for granted, than that he was considering something as striking as modal realism in Lewis' sense without explicitly pointing it out. Or, even more plausibly, we can say he was assuming for the sake of argument that some sense can be made of possible worlds talk, without committing himself to any particular interpretation. Be that as it may, the paper bears on our account of trans-world individuals.

I believe Quine's argument in (1976) can be summarized as follows. If we ask whether two momentary objects are stages of the same physical object, say a coin, then we can base our answer on considerations of "continuity of displacement, distortion, and chemical change" (1976, 861). On the other hand, if we ask whether something existing in another world is identical to the coin I have in my hand, such considerations do not apply. Predicates, such as 'coin' provide a "principle of individuation" in the cross-temporal case, but not in the trans-world case.

This is an interesting point. The charge is that, though we may be able to make *sense* of trans-world identity, the lack of principles of individuation makes *de re* modal talk vacuous: anything goes. Here is what I think the Quinean linguisticist should say in reply. In some worlds there will be many things that can be *reasonably* taken to be the manifestation of a possible individual x which has *this coin* as its actual manifestation. Some of these things, we may suppose, are F and some are not. If we want to know whether x is F in one of those worlds, we must make a decision as to which individual (if any) is its manifestation there. If all worlds turn out to be of this kind, then it will be reasonable to say that x is possibly F and it will be reasonable to say that x is not possibly F. Whether we say one thing or the other will be a matter of convenience or arbitrary choice. On the other hand, if we can find a world in which every reasonable manifestation of x is F, say a world in which everything reasonably taken to be my desk is red, then the only reasonable choice will be that my desk is possibly red. (Here I am assuming that we only say an object does not exist in a world if nothing can reasonably taken to be its manifestation there.) What determines if something counts as a reasonable manifestation is the *kind* of the object in question, together with general a priori modal principles about kinds. On this basis we could say, for instance, that any reasonable manifestation of this liter of water is made of H_2O molecules, that any reasonable manifestation of this desk is made of wood, etc. So I don't think it is entirely accurate to say that sortal predicates do not provide a principle of individuation in the trans-world case. It is rather that in this case the individuation works in a more roundabout way.

4. *Lewis' problem of reduction*

Section 3.2 of (Lewis 1986) remains the standard presentation and critical discussion of linguistic ersatzism. The linguistic ersatzist sees worlds as "complete, consistent novels"; but, as already mentioned, the notion of language in play is highly abstract and idealized. The "worldmaking language" doesn't need to be finitary and the words can be anything we please. In particular, anything can be its own name. In the end what makes linguistic ersatzism linguistic is just that the structures constructed represent ways the world could be *by stipulation*. Clearly, Quinean linguisticism is a form of linguistic ersatzism.

Lewis' first objection to linguistic ersatzism is that "modality must be taken as primitive" (1986, 150). This is what I called above the problem of reduction. Clearly, if modality must be taken as primitive, then the theory cannot be claimed to be reductive; and though a modal theory of possible worlds is not without interest, it is useless from an extensionalist point of view.

According to Lewis, the need to take modality as primitive comes in either via consistency or via implication. Via consistency, because the intuitive idea behind linguisticism is that a world is a *maximal consistent* set of sentences; and consistency here appears to be a modal, rather than a syntactic or semantic notion. Via implication, because sometimes it is the case that p according to a world w, not because w explicitly says that p, but because it says things that jointly *imply* that p; and implication here appears to be a modal, rather than a syntactic or a semantic notion. In both cases the challenge is to formulate the theory while avoiding the employment of modal notions of consistency and implication.

Take consistency first. Remember we said above that any felt incompatibilities between the predicates of the fundamental empirical theory could be incorporated into the definition of possible world. That arguably takes care of the problem of consistency, or at least of a good part of it. Certain basic combinations are ruled out syntactically, if need be. We don't need to worry about a world representing explicitly someone as being both married and a bachelor because 'married' and 'bachelor' are not part of the worldmaking language. We only need to take care, when it comes to defining means of implicit representation, to make sure these means are consistency preserving.

Of course, certain combinations are ruled out because we think they are impossible. But that is not a case of circularity. We are not analyzing the notion of possible world in terms of impossibility; but simply, as Lewis says, excercizing our understanding of the *analysandum* as we go. In the same spirit, if we choose a worldmaking language the predicates of which are, say 'occupied' and 'empty', that is because we believe everything supervenes of the distribution of matter over space-time. Different choices of vocabulary will reflect different metaphysical views. But again this is not a case of circularity.

Still, Lewis is not satisfied. He considers the case of a worldmaking language that represents point particles as being either positively charged or negatively charged. He then considers the idea that "we could declare outright that no ersatz world shall include two atomic sentences that predicate 'positive' and 'negative' of the same thing" (1986, 155). He objects that this way of proceeding is "risky". If we happen to be wrong and it is just a matter of (lawful) contingent fact that positive and negative charge never coexist, then by imposing this restriction we have falsified the facts of modality. On the other hand, if positive and negative charge really are incompatible, failing to impose the restriction falsifies the facts. Lewis adds: “If the ersatzist knew the facts of modality, he would know what to do. But he doesn’t and he can’t” (1986, 155). He concludes:

> The only safe course is to resort to primitive modality. The declaration must be conditional: *if* it is impossible for any one particle to be both positively and negatively charged, *then* let there be an axiom of unique charge. (Lewis 1986, 155)

I find this argument weak. Why can’t the ersatzist rely on his beliefs about the facts of modality and be content with it? Surely this involves a certain amount of risk; but the ersatzist might be willing to pay this price. Moreover, no one is entirely safe. For example, isn’t it risky to assume that, necessarily, for every way the world could be there is an isolated physical universe that is that way? Surely it is. And if this thesis fails, then the modal realist’s analysis of necessity in terms of concrete possible worlds also fails; and yet the modal realist is willing to take his chances. I suggest that the ersatzist is much safer than the modal realist, as far as reliance on modal knowledge is concerned.

Let’s turn now to implication and implicit representation. We want to say that possibly a donkey talks iff there is a possible world *according to which* a donkey talks. But what does that mean? If it means that there is a world such that a talking donkey *would* exist if the distribution of fundamental states over fundamental individuals were as that world says it is, i.e., if that world *strictly implies* the existence of a donkey, then we have circularity. What else can the ersatzist say? Let’s approach this problem by looking first at some easier cases of implicit representation.

If F is a basic predicate of the theory, and i is a basic individual, then for w to be a world according to which i is F, is just for <i, F> to belong to w. This is *explicit* representation. As a first easy case of implicit representation, we can consider truth-functional constructions. (In what follows, by ‘represents’ I shall mean ‘represents either explicitly or implicitly’.) A world w *implicitly represents* p as not being the case iff it does not represent p as being the case; and w *implicitly represents* p and q as being the case iff it represents p as being the case and q as being the case. In this way, taking atomic propositions as the base, we can recursively generate an account of implicit representation for all propositions truth-functionally constructed from atomic propositions.

Quantifications over basic individuals involving only basic predicates constitute a second easy case of implicit representation. It is the case that every basic individual is F according to w iff for every basic individual i, w contains <i, F>; and similarly for matrices truth-functionally constructed out of basic predicates, such as 'Fx or Gx'. (If we decide to think of basic individuals as existing contingently, then we would want to distinguish the proposition that every basic individual is F from the proposition that every existing basic individual is F. I will put aside this complication.)

Moreover, propositions about complex individuals could be easily represented, as long as they involve only basic predicates or predicates ultimately definable in terms of basic individuals and basic properties. For example, if F is a basic monadic predicate and R is a basic dyadic predicate, then we could define ' x is G' as 'x is the sum of two basic individuals y and z, and y is F and z is F and yRz'. Then we could say w implicitly represents it being the case that some complex individual is G iff there are two basic individuals x and y, such that w represents x as being F, y as being F and x as bearing R to y. (Here I am following Quine and assuming unrestricted mereological composition. If we had a more restricted view of composition, then worlds would also have to represent whatever relations are seen as generating sums. In any case we should have the part/whole relation as a basic relation of the language.)

Thus there is quite a lot that can be straightforwardly implicitly represented. For any statement A constructed from basic predicates via quantification over possible individuals and truth-functional operations we can say what it means for w to represent A as being the case. However, it is not clear that this is enough; and indeed most people think it is not enough. For it seems likely that we will not be able to define most of our predicates in terms of basic predicates, whatever the class of basic predicates turns out to be. Thus it is still not clear that we can say what it is for there to be, say talking donkeys according to a world.

But suppose for the sake of illustration that everything supervenes on the distribution of Democritean atoms. Then, vagueness aside, for every possible world, either there would be a talking donkey if things were, fundamentally speaking, as that world says they are, or there would not be a talking donkey. Call the class of worlds of the first kind T. Then, intuitively, there is a talking donkey according to a world w iff w belongs to T. In principle this can be turned into a definition. For to belong to T is to represent things as being fundamentally either like *this* or like *that* or …, where each disjunct is a full specification of which points are filled and which points are empty according to one of the worlds in T. Indeed, we can think of these disjuncts as being the worlds themselves. Thus, provided we believe ourselves to have found an adequate subvenience base, we should also believe in the existence of an extensional account of implicit representation.

The problem is that such an account would be "humanly impossible" to state, even given an adequate subvenience base. Consequently, if this is the best the Quinian linguisticist can do, there is a sense in which the theory cannot be completed and involves a very uncomfortable level of idealization. But we should not conclude that an extensional account of implicit representation is absolutely impossible. Also, an extensionalist may be happy to employ a non-extensional locution such as 'p according to w' provided there is *in principle* a way to define it in extensional terms.[10]

5. *Lewis' problem of expressive power*

Lewis' second objection is that linguisticists "cannot distinguish all the possibilities they should". This is, of course, the problem of expressive power I mentioned in the beginning. We have once more two cases. The first case is that of indiscernible possible individuals. Here is what Lewis says:

> Certainly it is at least possible that there should be many indiscernible individuals—alike in their intrinsic natures, and in their extrinsic properties as well. That would be so if there were two-way eternal recurrence. Or it would be so if the universe consisted of a perfect crystalline lattice, infinite in all directions. According to an ersatz world that represents such repetition in time or space, there are many indiscernible individuals. But we do not have correspondingly many indiscernible ersatz possible individuals, all actualized according to this ersatz world. One must do for all. What the ersatz world says, or implies, is that the one ersatz individual is actualized many times over. (Lewis 1986, 157f.)

This is supposed to happen because the many possibilities cannot be distinguished by the worldmaking language. In a world of two-way eternal recurrence, for example, there are many indistinguishable Napoleons, one in each epoch. So there seems to be infinitely many indistinguishable possible Napoleons. But according to Lewis there is only one ersatz indi-

10. There is also a problem concerning the representation of mathematical propositions. Here, given the infinitary character of the worldmaking language, one might appeal to the possibility of paraphrasing every mathematical truth as a tautology of a suitable infinitary language. See (Yablo 2002).

vidual, "only the one linguistic description of a filler of the role" (1986, 158). But I don't think that this is the case. The Napoleon of each epoch will be constituted by different basic individuals, by a different region of spacetime, as it may be. Or at least the temporal dimension will be different. Otherwise there would be no sense in saying they inhabit different epochs. So each Napoleon is distinguishable at least in terms of the epoch it inhabits. This seems to be implied by the very way in which the situation is described. *Pace* Lewis, it clearly doesn't matter that the recurrence is two-way, i.e. that the epochs are ordered like the integers and not like the natural numbers.[11] Then we will have the many possible individuals we want: each Napoleon can be seen as the manifestation of a different trans-world individual.

The second part of the problem of expressive power has to do with extra individuals and extra properties. The problem of extra individuals is the problem of representing possibilities for individuals that do not actually exist. Now it seems to me that the apparatus of trans-world individuals is capable of dealing quite well with these cases. On this view, a merely possible (trans-world) individual is a function from worlds to things (manifestations), with an empty manifestation at the actual world. Then to say, for instance, that there could be a cat different from all existing things is to say that there is a world w, such that, for some merely possible individual f, f(w) is a cat according to w. To take a more complex example, consider the proposition that, possibly, there is a talking donkey different from everything that actually exists, that is not famous but could be a famous. This is to say that some merely possible individual has a manifestation at possible world w which is an anonymous talking donkey according to w, and a manifestation at another world w′ which is famous according to w′.

As a last example, consider the more controversial case of worlds that differ only haeccetistically, i.e. over which individuals play which "roles". I see no reason to deny the existence of such possibilities. Taking the Democritean model as an example, we can imagine a world in which everything happens exactly as it does, only five minutes earlier. Then we would have different spacetime points instantiating the same basic properties and relations. Different examples could be constructed for more complicated versions of the view. But what should we say about such possibilities as the possibility that Adam and Noah swap roles? Here we must

11. See (Lewis 1986, 63).

think about Adam and Noah as trans-world individuals, of course. We can then accommodate this possibility by imagining a world in which Adam's manifestation plays the same role played by Noah's actual manifestation and vice-versa. Such worlds could still differ: one could, for instance, be a world in which everything happens fiver minutes earlier.

It is quite otherwise when it comes to things. Here the Quinean simply denies the possibility of extra ones. But to do this is not to fly in the face of our pre-theoretic modal intuitions. We have no pre-theoretic intuitions about fundamental individuals and their sums, for the simple reason that they are highly theoretical entities. It would be strange indeed if we denied the possibility of there being a rat in my studio right now, or the possibility of there being something distinct from anything that actually exists. But we are not denying these claims, which can be easily dealt with using our apparatus of trans-world individuals.

Finally, let's look briefly at the much discussed problem of "alien properties". Here is what Lewis says:

> I complain that if we only have words for the natural properties that are instantiated within our actual world, then we are not in a position to describe completely any possibility in which there are extra natural properties, alien to actuality. (Lewis 1986, 159)

One could choose to follow Armstrong (1989) and simply deny the possibility of alien properties. I don't think this is completely out of the question. But maybe the Quinean has the resources to account for them. Of course, the Quinean does not believe in the existence of properties, alien or otherwise, natural or otherwise, prior to their construction via possible worlds. For the Quinean linguisticist, properties of individuals are functions from possible worlds to sets of things. Thus at each world a property has an extension. A trans-world individual is said to have a property at a world iff the thing that is its manifestation at that world is in the extension of the property at that world. Some of these properties will not be instantiated in the actual world. Call these properties *alien properties*. An alien property is an alien *natural* property in a world w if it is natural in w. But what is it for a property to be a natural property of a world?

Well, maybe the distinction between natural and non-natural properties is basic. In any case, if the modal realist can take the distinction as basic, as Lewis thinks he can, then I don't see why the same procedure would not be available to the Quinean. Alternatively, one might attempt to define the notion of a natural property. Suppose one wishes to define it in terms of laws: the natural properties are those figuring in laws of nature. In this case the Quinean could say that whether or not a property will figure in a law is determined by its pattern of instantiation. Similarly for other ways in which one might attempt to define the notion of natural property. What the Quinean clearly *cannot* accept is that there are properties such that two worlds may differ in relation to it while at the same time being identical in instantiation of fundamental properties. For that amounts to acknowledging that worlds do not determine entirely what is the case, which is contrary to the notion of a world. Alien properties in this sense must simply be denied.

This is, then, how I suggest the Quinean should deal with the problem of expressive power. But I should mention that there are other conceivable ways. These other ways involve enriching the worldmaking language so as to be able to express more. Towards the end of his discussion of linguistic ersatzism Lewis makes a suggestion in this direction, although he still didn't find the result satisfactory. More recently, (Heller 1998, Melia 2001, and Sider 2002) are interesting contributions to this approach. Though they are not themselves Quinean linguisticists, their proposals can be easily adapted. Sider's idea of replacing talk of worlds by talk of the whole *pluriverse*, so as to be able to bind variables across possible worlds and thus represent the same alien property or individual existing in different worlds, seems especially clear and promising. Of, course, if the preceding discussion is along the right lines, then such expansions are not needed.

6. *Final remarks*

As we saw, Quinean linguisticism faces a number of serious challenges. This is actually not surprising, given the strictures of extensionalism. But in fact, as (Divers 2002) makes clear, many less restrictive views face analogous objections when attempting to give a reductive account of modality and of possibilist ontology. They are better off only in that, not being forms of extensionalism, they are not faced with the choice between reduction and elimination.

What we must do is consider the pros and cons of each alternative theory and make our choice. If one decides to drop extensionalism, then one is free to adopt a *modalist* position, in which some modal notion is taken as primitive. But if one is a "confirmed extensionalist", then the only alternatives to Quinean linguisticism (or some variant of the same basic idea) appear to be modal realism and eliminativism; and given *that* range of options, I think Quinean linguisticism still has the upper hand, in spite of its difficulties. In particular, it seems clear that no matter how many bullets the Quinean linguisticist has to bite, the result will still be more congenial to our intuitions than the eliminativist's declaration that modal discourse makes no objective sense and that intensional ontology must be entirely forsaken. Someone who, like Quine, takes extensionalism more or less as a fixed point should be a Quinean linguisticist.

References

Armstrong, David 1989: *A Combinatorial Theory of Possibility*. Cambridge: Cambridge University Press.

Bigelow, John C. 1988: "Real Possibilities". *Philosophical Studies* 53, 37–64.

Carnap, Rudolf 1956: *Meaning and Necessity*. Chicago and London: University of Chicago Press.

Cresswell, Max J. 1972: "The World Is Everything that Is the Case". *Australasian Journal of Philosophy* 50, 1–13. Reprinted in: Michael Loux (ed.) 1979, *The Possible and the Actual*. Ithaca: Cornell University Press.

Divers, John 2002: *Possible Worlds*. London and New York: Routledge.

Heller, Mark 1998: "Property Counterpart in Ersatz Worlds". *The Journal of Philosophy* 95, 293–316.

Lewis, David 1986: *On the Plurality of Worlds*. Oxford: Basil Blackwell.

Melia, Joseph 2001: "Reducing Possibilities to Language". *Analysis* 61, 19–29.

Prior, Arthur N. & Fine, Kit 1997: *Worlds, Times and Selves*. Cambridge, MA: University of Massachusetts Press.

Quine, Willard V. O. 1950: "Identity, Ostension, Hypostasis". *The Journal of Philosophy* 47, 621–633. Reprinted in: Willard V. O. Quine 1980, *From a Logical Point of View*. 2nd revised edition, New York: Harvard University Press.

— 1960: *Word and Object*. Cambridge, MA: The MIT Press.

— 1968: "Propositional Objects". *Crítica: Revista Hispanoamericana de Filosofía* 2, 3-29. Reprinted in: Willard V. O. Quine 1969, *Ontological Relativity & Other Essays.* New York: Columbia University Press.

— 1976: "Worlds Away". *The Journal of Philosophy* 73, 859-863. Reprinted in: Willard V. O. Quine 1981, *Things and Theories*. New York: Harvard University Press.

Quine, Willard V. O. 1976a: “Whither Physical Objects”. *Boston Studies in The Philosophy of Science* 39, 497–504.
— 1981: *Things and Theories*. New York: Harvard University Press.
— 2001: “Confessions of a Confirmed Extensionalist”. In: Juliet Floyd & Sanford Shieh (eds.), *Future Pasts: The Analytic Tradition in Twentieth Century Philosophy*. New York: Oxford University Press, 215–222. Reprinted in: Dagfinn Føllesdal & Douglas B. Quine (eds.), 2008, *Confessions of a Confirmed Extensionalist And Other Essays*. Cambridge: Harvard University Press.
Sider, Theodore 2002: “The Ersatz Pluriverse”. *The Journal of Philosophy* 99, 279–315.
Sklar, Lawrence 1977: *Space, Time and Spacetime*. Los Angeles: University of California Press.
Skyms, Brian 1981: “Tractarian Nominalism”. *Philosophical Studies* 40, 199–206.
Smullyan, Arthur F. 1948: “Modality and Description”. *The Journal of Symbolic Logic* 13, 31–37.
Tarski, Alfred 1956: “The Concept of Truth in Formalized Languages” In: John Corcoran (ed.), *Logic, Semantics, Metamathematics*. 2nd edition, Indianapolis: Hackett Publishing Co.
Yablo, Stephen 2002: “Abstract Objects: A Case Study”. *Noûs* 36, 220–240.

Grazer Philosophische Studien
90 (2014), 229–230.

RUDOLF HALLER (1929–2014)

Rudolf Haller, who founded *Grazer Philosophische Studien* in 1975, died after a long illness on February 14, 2014. From 1975 to 2000, he published 59 volumes of this journal and secured its high international reputation. The dedication and energy with which he worked for the *Grazer Philosophische Studien* were extraordinary and are inspiring still.

Rudolf Haller studied philosophy, philosophical sociology, history, and history of art at the Universities of Graz and Oxford. From 1961 to 1967 he was teaching philosophy in Graz and at the Pädagogische Hochschule in Hannover, before taking up the position of Professor of Philosophy at the University of Graz, which he held until 1997. From 1982 to 1997 he also directed a research institution for Austrian Philosophy that he had founded (*Forschungsstelle und Dokumentationszentrum für Österreichische Philosophie*).

His publications are numerous (5 books and more than 400 articles) and cover a wide range of topics both historically and systematically. One of the areas where his work will have a lasting influence on future generations of philosophers was his research on, and unremitting commitment to, Austrian philosophy, in particular the philosophy of Ludwig Wittgenstein, the Vienna Circle, the Brentano School, and the Graz School (Alexius Meinong and his students). He co-edited the writings of Otto Neurath, and he was the chief editor of the *Alexius Meinong Gesamtausgabe*. Among his many organisational achievements, one should mention that in 1976 he was among the founders of the Austrian Ludwig Wittgenstein Society and one of the initiators of the annual Ludwig Wittgenstein Conferences. In 1979 he initiated a book series dedicated specifically to the study of Austrian Philosophy and contributed two volumes to this series: *Studien zur Österreichischen Philosophie*: *Variationen über ein Thema* (1979), and *Fragen zu Wittgenstein* (1988, engl. transl. *Questions on Wittgenstein*, Routledge 2014).

As his successors in editing the *Grazer Philosophische Studien*, we would like to mention here specifically his contributions as an author to this journal. Apart from several introductions to special-topic volumes and letter editions, he contributed no less than 11 papers and 8 reviews. Even if these contributions are not representative of his entire work, they may serve to illustrate the breadth of Rudolf Haller's intellect.

The papers with a historical topic clearly reflect his main interests in this era. Four of them are about Ludwig Wittgenstein (vols. 21, 33/34, 42 and 49), two are about Franz Brentano and his school (vols. 5 and vol. 24), one is on Alexius Meinong (vol. 50), and one is on Moritz Schlick and Otto Neurath (vol. 16/17). The three papers with a systematic focus are extremely diverse. One of them is on relativism (vol. 44), while the other two do not easily fit into a single philosophical discipline. They are partly about matters in ontology, and partly located at the borderline between aesthetics and philosophy of science (vol. 12/13, vol. 20).

With these few examples in mind, it appears that Haller himself described the character of his philosophical work quite well when he once remarked about his early writings: "They remind me that unity in diversity has not only been regarded as the essence of art, but is also a maxime of the reflection and self-reflection of thought in general." ("Sie erinnern mich daran, dass die Einheit in der Mannigfaltigkeit nicht nur zum Wesen der Kunst gerechnet wurde, sondern auch eine Maxime der Betrachtung und Selbstbetrachtung des Denkens überhaupt bildet. (*Facta und Ficta*. Stuttgart 1986, 7).

We would like to thank Rudolf Haller for all he has done for philosophy in general and for this journal in particular. As a final gift Rudolf left us, we found a short autobiographical sketch (probably composed in 1995), in which the founder of this journal relates, and reflects on, his own academic life. We would like to share it with our readers. It is our honor to publish it here.

The Editors

Grazer Philosophische Studien
90 (2014), 231–244.

AN AUTOBIOGRAPHICAL OUTLINE

Rudolf HALLER

The historical background of my schooldays was the last period of the Austrian Nazi regime and the years when Austria no longer existed after it had become part of the German Reich. After the war, I was sent to school to Graz, and I took my final exams in 1947/48 at a night school.

How I came to read Nietzsche, Hölderlin and Vaihinger (!) at the age of fourteen, I do not remember. But I remember well that at the same time I began to develop a philosophical theory of the world, to shape my unclear thoughts, imitating the style of Hyperion and Zarathustra. My idea of the world, i.e., the entire cosmos, was not at all original. I saw it in the most perfect infinite spherical form, consisting of material and immaterial parts, altogether of a divine nature.

What induced me one night to destroy all my previous "works" (including a Leibniz-like monadology from my boyhood, poems, and letters I had received) can only be explained by my experience of a principal inner change. Following my inclination to indulge in religious feelings that had become more and more important to me, I thought it necessary to destroy the past, i.e., my earlier literary productions. However, I decided not to give up philosophy, although I then did not know very well what that was. The Austrian grammar-school curriculum provides the subjects psychology and philosophy in the final two years. My intention to be true to philosophy was first fostered by my philosophy teacher Maria van Briessen, and later by the bookseller and publisher Filip Schmidt-Dengler, whom I first met in his second-hand bookshop. In the late thirties, his financial situation had permitted him to establish a small publishing house. Among other things, he wanted to offer to the general reader works by the Austrian disciples of the poet Stephan George, to whom he belonged himself under the pseudonym Filip Rabus.

I started to study philosophy at the University of Graz in the winter term of 1947/48, when there were only about 3000 students overall. Many of the students were much older than the usual beginners, because they had served in the war or had been prisoners of war. That made our student-years much different from the ones of the others who followed. The students were working hard and were seriously trying to do their best to finish their studies within the shortest possible time, because many of them had to study under difficult economic conditions or wanted to make up for their wasted years in the war. As the Austrian university system is not structured in classes, beginners were attending lectures together with students who were already preparing their doctoral theses. There was another difference from the American system, regarding university degrees: at the Philosophical Faculty there was no other degree than the Ph.D.; and only those students who had qualified for being grammar-school teachers received the title of professor.

There were only two teachers at the Philosophical Seminar, as the Philosophy Department was then called: Konstantin Radakovic and Amadeo Silva-Tarouca. Konstantin Radakovic, a brother of the Vienna-Circle member and mathematics lecturer Theodor Radakovic, was the only faculty member who had spontaneously given up his post after the invasion of Nazi troops into Austria in 1938. His special subjects were history of philosophy and philosophical sociology. Each semester he gave courses on a topic from the history of philosophy, ranging from the pre-Socratic period to the 19th century. He held an empiricist point-of-view, actually a Humean one, combined with an open-mindedness towards attitudes that differed from his own. Neither in his lectures nor in his critiques did he make the slightest effort to convert his students or to convince them of his own ideas. His mild scepticism became apparent in his seminar, where he often left it to the students to choose a topic. The fact that, on the occasion of his 65th birthday, he was honoured by his students with a *Festschrift*, entitled *Philosophie der Toleranz* (1960), resulted from his generous attitude towards people who developed theories and theses opposite to his own.

The second philosophical chair was held by Amadeo Silva-Tarouca. Whereas Radakovic had published only a small number of minor writings but had been a university teacher for many years before he was offered the chair, Silva-Tarouca, descending from a well-known aristocratic family, had published quite a number of books on various topics of which I just mention *Weltgeschichte des Geistes* (1939), *Deutsche Kunst aus deutscher Vergangenheit* (1943), and *Thomas heute* (1947), the last being a kind of existentialist interpretation of Thomas of Aquinas, indicating better than the other titles the philosophical direction which later became manifest in his *Philosophie der Polarität* (1955). In the course of eight semesters, Silva-Tarouca developed his own system of philosophy, called "Ontophenomenology". It was supposed to be able to unify antagonistic poles at a higher level, such as, for instance, "for me/without me", "subjectivity/objectivity", "thinking and wanting", in a polar-dialectic manner. Since I had to do the work of Silva-Tarouca's scientific aid, much time was spent discussing his invention of this system, whose terminology was strongly influenced by existentialism. At that time, every doctoral student of the Philosophical Faculty was obliged to sit for a one-hour oral exam after having submitted their doctoral theses, i.e., half an hour with Radakovic and Silva-Tarouca each. As there was no secretary yet at the Philosophical Seminar, the scientific aid was a secretary, librarian, personal assistant to the professor, and tutor at the same time.

The two professors had one assistant by the name of Rudolf Freundlich and a parttime scientific aid. The latter job was first done by Dr. Pfniß, later by Georg Janoska, and from 1954 on by myself. Freundlich, who had studied with Moritz Schlick and Karl Bühler and whose dissertation had been supervised by Robert Reininger, qualified as a university lecturer in 1948. He lectured on present-day philosophy and on logic and philosophical logic, in a series of lectures, covering first the phenomenological school and German existential philosophy, and, in a second series, logical empiricism and modern ontology. In the logic course, students for the first time got acquainted with the principles of mathematical logic, which five of us had already been taught in an optional course by the professor of mathematics, Hermann Wendelin. Maybe at that time, or earlier, Rudolf von Scheuer drew my attention to Wittgenstein's *Tractatus*, which I eagerly made extracts from, because the dichotomy between saying and showing as well as his "silence proposition" seemed to me particularly helpful in my religious reflections of those days.

In my fourth year as a student, it became necessary to choose a dissertation topic. Since my religious thoughts seemed to me the most interesting, I chose an interpretation of the religious existentialism of the Russian Jewish philosopher Leo Shestov. He had been suggested to me by Filip Schmidt-Dengler, who had published a German version of Shestov's main work, entitled *Athen und Jerusalem: Versuch einer religiösen Philosophie* (German transl., Graz 1938). Shestov, a radical anti-rationalist, vehemently attacks the traditional Greek-Latin image of God and the traditional justification of religious belief. He opposes all rational explanations of the Bible, taking it literally and denying any metaphoric or symbolic interpretation of the text. It was mostly this uncompromising radicalism in the spirit of Kierkegaard's "Either-Or"-attitude that attracted me. One of Shestov's later works, which also appeared in German, was dedicated to the Danish philosopher: *Kierkegaard und die Existenzphilosophie*. Kierkegaard's radical criticism of all efforts to rationalize religious belief, and Christian belief in particular, did not convince me, however, and I much favoured Shestov's alternative: either the works of God are governed by the rules of logic and are not contrary to reason, in which case many of them are false and unacceptable, or God's almighty power as well as his works are not subject to, or restricted by, logical rules or the laws of nature, and it remains true that to believe means to lose one's brains to regain God instead—nowadays one would call this a fundamentalist thesis. According to Shestov, the purpose of philosophy is not mere reflexion, but struggling for the one thing that is necessary. One example, frequently used by Shestov to demonstrate and defend his conception of belief, is Abraham who obeyed God's command to leave his country without knowing where to go. It was this unquestioning belief that strongly attracted me and also brought about a solution to the very subject of my dissertation, a solution that had actually been found by Kierkegaard. In the event, the completion of my dissertation put an end to my religious excursions, leading me towards conceptual and aesthetic questions. Accordingly, the first lecture I ever delivered was not on Shestov, but on Nietzsche and Sartre. The shift in topic that this change involved reflected my rising interest in existentialism, which, apart from France, had come into fashion in Central Europe.

I have always been interested in aesthetics and the philosophy of art, and I noticed very early that dealing with these topics belongs to the most difficult things in philosophy. Except for my first two semesters as a university teacher and a later seminar on Nelson Goodman's *Languages of Art*, I did not use topics from aesthetics for my lectures again. My interest in this difficult specialty has remained, however, even though I have published very little on it. The natural cause for my theoretical interest in aesthetics must have been the power of the arts, or better, the works of art that attracted me strongly. As far as I remember, that has always been the case and relates to nearly all kinds of poetry, music, painting, pictorial art, and architecture. It is not at all surprising that many of my friends have been writers, artists, and architects, as well as philosophers and scientists. Since my early student days, among my friends have been the architect and critic Rudolf von Scheuer, the poets Heimrad Bäcker, Rudolf Stibil, and Alfred Kolleritsch, the painters Mario Decleva and Hannes

Schwarz, and the pianist Alfred Brendel, who at that time also painted and wrote poems.

One of my most interesting fellow students was Georg Janoska, who wrote his dissertation on *Ding an sich und wissenschaftliche Philosophie* and had completed his studies two years before myself. Thus, for a few years the confrontation between Kant's theory of belief and Shestov's conceptions was one of the subjects of our conversations. At that period, I learned a lot from my friend Janoska; we shared several philosophical ideas, and, additionally, the topics he was mainly dealing with, like Kant, Hegel, Marx, and Freud, had not been familiar to me before. Janoska favoured Franz Kröner's so-called "Systematology", a kind of metaphilosophy of philosophical systems, that, without occupying any individual point-of-view, was to lead to a systematic judgment of the "anarchy of philosophical systems". Kröner, who later sympathized with the Nazis, published a book in 1929 under that title (*Die Anarchie der philosophischen Systeme*), trying to prove the idea of a necessary pluralism of philosophical systems, deriving from the very nature of system- or theory-building itself. However, this semi-sceptic position of an anarchy of philosophical systems—as it had earlier been called by Dilthey, who vehemently opposed this thesis—was not to be taken as relativism. Just as pluralism of different geometries does not imply relativism, so different alternatives or antitheses, inherent in any theory or system, do not back up universal relativism of philosophical theories. Thus, the theorem about the parallels in the Euclidian geometry holds unrestrictedly true, regardless of the fact that there are contradictory theorems in other geometries.[1]

Of course, we raised the question of theory pluralism too early, especially concerning philosophy, maybe because we had no firm standpoint then and no analytic comprehension of the nature of theories and their changes. Later this became a favourite subject of my lectures and seminars after my return to Graz from Hannover. In the following years, Janoska's manifold interests more and more led him to the history of philosophy, ranging from Kant to Hegel and Marx; eventually he concentrated on something that he called "nominalist dialectics", a topic I could not and would not make my own.

Reading the philosophical literature that was available in the mid fifties, I came to the conclusion that for continuing my own work I should have

1. Franz Kröner, *Die Anarchie der philosophischen Systeme* (1929). Geleitwort zur Neuausgabe von Ferdinand Gonseth; Nachwort von Georg Janoska. Graz: Akademische Druck- und Verlagsgesellschaft 1970, pp. 327.

to know more about analytic philosophy. With the aid of my friend Helmut Sihler, later general manager of the Henkel company in Germany, I applied for a post-doc scholarship to as many as 42 American philosophy departments. There were no offers, however, except from the Universities of Chicago and Yale, but with restrictive financial conditions. I was fortunate enough to be offered an Austrian scholarship at Oxford, where in 1958/59 I received my "philosophical baptism". I studied under Gilbert Ryle and shared his opinion about the importance of Wittgenstein's *Philosophical Investigations*. I found myself in the Eldorado of philosophy, which Oxford then undoubtedly was. J. L. Austin had just returned from his successful stay in the United States. He lectured on "Sense and Sensibilia", in 1959, for the last time, and on "Words and Things" in the previous term, or the following one. I participated in David Pears and Brian McGuinness' "Introduction to the *Tractatus*", had good conversations with A. Quinton, listened to Peter Strawson's lectures, and to Isaiah Berlin's inaugural lecture, and got acquainted with Friedrich Waismann, who, though being frequently ill, had still remained interested in what was going on. I enjoyed taking part in evening lectures and discussions of various societies, as for instance the Undergraduate's Jowett Society, the Socratic Club, and the Philosophical Society, which were as enthusiastically attended by professors and lecturers as by visitors. I still have a vivid memory of the whole philosophical *furioso*, and of course of Gilbert Ryle, who patiently listened to my stammering English and tried to justify his philosophizing against the background of his early studies of Meinong, Husserl and Heidegger. His first-rate seminar on "Late Plato and early Aristotle", with participants like Elizabeth Anscombe, Richard Walzer, G. E. I. Owen, John Ackrill, Brian McGuinness, and others, was highly impressive. Ryle was marvellous in expounding the reason why a problem stated by Plato was a real problem or not, and he returned to Plato's *Parmenides* again and again as well as to Aristotle's *Categories*, topics with which Ryle had grappled already in the late thirties. His "informal instructions", held weekly as open discussions, were an excellent example of his way of taking up philosophical problems. The audience liked this way of doing philosophy, and so did I. What later came to be called "linguistic philosophy" was in fact nothing else but a form of discussing philosophical problems in a serious manner, paying careful attention to what was said in class, and analyzing it as far as our linguistic skills allowed us to.

Returning from Oxford, I started to write my habilitation thesis: *Untersuchungen zum Bedeutungsproblem in der antiken und mittelalterlichen Phi-*

losophie[2] (*Investigations of the Problem of Meaning in Ancient and Medieval Philosophy*), which would give me the right to lecture at the university. During my habilitation colloquium in the academic year 1960/61, attended by all professors of the whole faculty, which then included humanities and natural-science departments, everybody was entitled to ask me any questions from their specific perspectives. To give an example, the professor of classics and Byzantine studies, Endre von Ivanka, examined me on Scotus Eriugena's conception of universals, on Descartes, Goethe, and finally on Sartre, the latter being not at all difficult for me.

The history of concepts and ideas is quite fascinating, like any investigation of historical changes. But without a philosophical point-of-view it does not contribute to any systematic construction. However, it can be helpful in viewing philosophical problems in a new perspective. I had got an idea what meanings really were and thus was able to grasp what others proposed as their theories of meaning and reference. Very often we discover that our philosophical ancestors found solutions that are valuable and useful even today. This holds true for myself, when in later years I spent a lot of time excavating the forgotten history of Austrian philosophy. To understand what kind of questions can be asked is an important step in our philosophical practice. But I did not find out all that by myself. Ryle was one of those from whom I learned the most, another one was Roderick M. Chisholm.

After Oxford philosophy had been revealed to me, I returned to Graz, where I was asked by Hofrat Kindinger whether we could host the Chisholm family in the academic year 1959/60. We invited him to our house, although there was not enough room for visitors at that time, only six rooms altogether, without central heating. So Roderick Chisholm worked hard to heat three rooms and the kitchen. But as he remarked in his "Self-Profile", it "meant the beginning of a lasting friendship and a series of mutual projects pertaining to philosophy in Austria".[3] In 1972, Chisholm was awarded an honorary doctoral degree by our university, and from 1974 until 1991 he regularly gave so-called "Blockseminare" in Graz, four or five weeks after the end of his spring term at Brown. He has been honorary professor at our department for seventeen years, and his impact on myself, the assistants, and the students has been remarkable and lasting.

2. Published in: *Archiv für Begriffsgeschichte* 7 (1962), 57–119.
3. R.M. Chisholm, "Self Profile", in R. Bogdan, ed., *Roderick Chisholm*, Reidel 1985, 11.

But now I should continue sketching my own career. After having qualified as a university lecturer in 1961, nearly everything I was working on was determined by analytic literature, and I obtained a new perspective that induced me to see traditional philosophy in a new light and to follow the tradition of Austrian philosophy, based on the views of Russell and G. E. Moore and their successors. There was a general feeling as if the spirit of the world had moved from the old continent to Great Britain and over to the new continent, the United States, because it rather seemed as if, after the heyday of Oxford philosophy, the new analytic ideas came mainly from the United States.

I started lecturing in the winter term 1961/62 on "Descartes and Modern Subjectivism", gave a pro-seminar course on "Kant's Critique of the Aesthetic Judgment", and continued with lectures on Leibniz. In 1962, I gave my first seminar on "Wittgenstein and Ryle". Other topics I was concerned with were "Theory of Meaning", "Analyticity", "Perceiving and Knowing", and "The Concept of a Person". In aesthetics, where I had contributed a paper on "The Present Situation of Aesthetics" at the XII. International Congress of Philosophy in Venice in 1959, I insisted on taking into account two pillars: the work of art as object and aesthetic experience. Therefore, the analysis of the aesthetic object has to comprise both sides. As the methods of analysis I proposed psychological, phenomenological, and semiotic ones.

In 1962, I delivered a lecture on the problem of meaning at the German Congress of Philosophy in Münster, which provoked a lively discussion with P. Lorenzen, G. Patzig, H. A. Schmidt, etc., and in 1963, I was invited by Hans Wagner to the University of Bonn, where I spoke about "The Linguistic Method in Philosophy" and received a rewarding response. H. Wagner, E. Rothacker, G. Martin, G. Hasenjäger, Baron and Derbolav, among others, were taking part in the discussion. Especially Hans Wagner, who was very open-minded concerning the analytic movement and with whom I had a good rapport, supported me a great deal for some time. From 1963–65, I also contributed to Paul Weingartner's "Forschungsgespräche" in Salzburg, with papers on description theory, the controversy over the analytic-synthetic dichotomy, and on metaphysics and language. Especially the second of those symposia was of lasting influence on myself. There I met Herbert Feigl, Paul Feyerabend, Jaakko Hintikka, Bela von Juhos, and Werner Leinfellner. Hintikka delivered a few lectures on analyticity, and we had very good and profound discussions. At that period I learned to admire Quine's analysis and followed Putnam's advice: "Ignore the analytic-synthetic distinction, and you will not be wrong in connection with any philosophical issues not having to do specifically with this distinction."

In 1964, I received the title of associate professor, which did not mean very much. In 1965, I was invited for the first time as a visiting professor to Munich to stand in for Professor Stegmüller during his sabbatical. The audience there was well prepared to hear my stories about Oxford philosophy. One of my students was Eike von Savigny, who a few years later published an account of the "philosophy of normal language" (*Die Philosophie der normalen Sprache*, 1969), which to a certain extent continued to expound the ideas of ordinary language philosophy, a topic I had dealt with in my seminar. In the winter term, I proposed to lecture on "Perceiving and Knowing". Starting from the difference between sensing and perceiving, I discussed some of the proposals of Austin, Ayer, Chisholm, Malcolm, Ryle, and Wittgenstein, trying to find my own line. Particularly interesting seemed to me the significance of our memory in all our language use and in our mental habits. The problem of the adequate analysis of remembering sounds, tunes, colours, etc., can best be solved by referring to existing objects. Soon after the term in Munich, during a further term in Oxford, supported by a grant from the British Council, I accepted a chair at the Pädagogische Hochschule in Hannover.

I have always been interested in putting ideas into practice through institutional organizations, such as societies, symposia, congresses, and editions that I have been able to establish in the course of the years, an effort that has taken up much of my time, sometimes too much. Even to those institutions that had come into being without my initiative, I devoted more energy than was good for me and my family.

I founded the Shestov Society together with Schmidt-Dengler, as well as an assistants association, the "Vereinigung für wissenschaftliche Grundlagenforschung", in 1964, which, within the first twenty years of its existence, organized nearly a hundred lectures, mainly delivered by foreign philosophers and theorists of science. It was not until 1975 that the idea of a project for an international journal of analytic philosophy had successfully been carried out, thanks to the Dutch publisher Schipers and later to his son, Fred van der Zee. I quite deliberately called it *Grazer Philosophische Studien*, a journal now holding a special position in the history of philosophy. It is presently producing its 50th volume, which contains all the contributions to the Meinong Symposium of 1995. In 1979, I started a further series with Rodopi Publishers, named after my book *Studien zur Österreichischen Philosophie*, to materialize another idea of mine that was of great importance to me. Up to now, more than 25 volumes have been published.

I should like to mention two further large-scale projects. One was on "Science and Ethics", initiated by Ivan Supek, now President of the Croatian Academy of Sciences, on the occasion of my participation in an inter-university program in Dubrovnik in the early sixties, sponsored by the German Volkswagen Foundation. There was an impressive list of outstanding participants from Finland, Sweden, Norway, Germany, France, Great Britain, Poland, Yugoslavia, and the USA. Although we rather discussed about science than about ethics, there had been made a start anyway. In the course of nearly ten years, the contributions to that project resulted in three volumes, edited by Risto Hilpinen, Keith Lehrer, and myself.[4] The other project I should mention is my cooperation with the Austrian Wittgenstein Society which started to organize a series of symposia in Kirchberg with the help of the veterinarian A. Hübner, his wife Lore, and Paul Weingartner, and has become a centre of lively Wittgenstein discussions.

Perhaps it is also important to mention that, after having accepted the call to the University of Graz, I organized exchange programs with several American philosophy departments, which enabled lecturers from Boulder, Minneapolis and Tucson to teach in Graz and, vice versa, enabled our assistants to teach in the USA. I will now stop listing my activities—activities I do not regret but, on the contrary, have always felt necessary to take part in, according to my conviction to do what had to be done.

During my stay at Oxford, I was offered a chair in Hannover to succeed Gustav Heckmann, a devoted student of Leonard Nelson.

A few weeks after I had begun to teach in Hannover, the capital of Lower Saxony, I rather surprisingly received a call to Innsbruck. As soon as that had become known in Graz, I was suggested for a chair there. Thus, within a few months, I was offered three chairs, but unhesitatingly decided to accept the Graz offer, because the conditions, like for instance the size of the department, its library, and my private circumstances, seemed to me most convenient. So the "lifetime professor" of Hannover stayed for only one year in Hannover, where I much enjoyed teaching, because the students were not as shy as in Austria and were keen on discussing everything that had been said and also answered eagerly, when asked at examinations.

4. R. Hilpinen, ed., *Rationality in Science: Studies in the Foundation of Science and Ethics*. Dordrecht-Boston 1980; R. Haller, ed., *Science and Ethics*. Amsterdam 1981; K. Lehrer, ed., *Science and Ethics*. Amsterdam 1987.

My time in Hannover was the only one in my life that I happened to get in contact with a group of young actors and stage directors with whom I spent a lot of time. After having accepted the chair in Graz, I continued teaching in Hannover for one semester. Even today I still remember with pleasure the marvellous period in Hannover.

To teach in Graz meant to start from the very beginning. There were two other chairs here, the one held by Rudolf Freundlich, and the other one held by Amadeo Silva-Tarouca, who was about to retire and was followed by Ernst Topitsch, who, like myself, had taught in Germany for a while before receiving a chair in his home country. While Freundlich mainly taught on logic and language theory, Topitsch was responsible for the history of philosophy and periodically offered classes thereon.

I set out to design a cycle of lectures for a period of four semesters, exclusively dealing with philosophy of science, a discipline I had gone further into six years before. What was closer at hand than to continue the Austrian tradition of the philosophy of science, i.e., the Vienna Circle philosophy, and particularly Stegmüller's elaborate presentation of *Probleme und Resultate der Wissenschaftstheorie*, published no earlier than two years after I had commenced teaching in Austria again. At that period, philosophy of science had reached a climax, and it was important and interesting to elaborate the Carnap-Quine-Putnam line. Equally useful to me was Mario Bunge's publication *Scientific Research* (1967), but even more appealing I found—apart from Quine's writings—the publications by Carl G. Hempel, ranging from *Fundamentals of Concept Formation in Empirical Science* (1952), to *Aspects of Scientific Explanation* (1965) and *Philosophy of Natural Science* (1966). I was much inclined not to disregard the historical origin of phenomena. Feyerabend's and Kuhn's publications, also available in German of late, presented a new interpretation, forming a theoretical background for Ernst Mach's conviction that history has made everything and history can change everything. From reading Neurath, which Kuhn and Feyerabend certainly also had done, I knew what role the scientific community was able to play.

In general, I can say about that time of a new start, that it was my main concern to bring information and cooperation to the same level as quickly as possible, which had not been the case between Austria and the leading countries in analytic philosophy. The first steps should be to freshen up the stuffy, provincial atmosphere, and to immediately recommend to the more gifted students to go abroad for one or two semesters, for instance, to one of the well-reputed German universities, such as Göttingen, where Patzig, Scheibe and Wieland were teaching. In addition, I planned symposia and discussions in order to exchange ideas and opinions and to point out the relation between different points of view. Philosophers were invited from Scandinavia, Germany, the Netherlands, England and the USA, and it was most delightful to see that all members of the department and the students were cooperating enthusiastically, and many of them followed my advice to go to Göttingen or to England for a while. There was a high-spirited atmosphere then, which I cannot imagine any longer in the present days.

The next subject I started to work on was the theory of knowledge, and I translated Chisholm's *Theory of Knowledge* (2nd edition, 1977) into German. I agreed with him on various points but found that the method of doubting should be given more attention, and that justifying knowledge could most adequately be carried out by acting in analogy with performative speech acts. It was the analysis of Wittgenstein's later writings which led me to the thesis of praxeological foundationalism in the theory of knowledge. How much I owe to Wittgenstein, whom I regard as the most outstanding philosopher of this century, cannot be put into words. When working on him, I have always found new questions, and there has not been any other philosopher whom I have been dealing with more intensely.

My most important duty as a university teacher has been to provide a profound education and training for my students that should keep up with international standards, a goal that can be reached, for instance, by inviting well-reputed foreign philosophers who directly demonstrate to students how to make good philosophy. So I regularly invited Roderick Chisholm (Brown University) and Stephan Körner (Bristol University), who both were awarded honorary degrees by our university. A few years later we had two other visiting professors: Keith Lehrer (University of Arizona, Tucson) and Brian McGuinness (Oxford and Siena). Every year I have tried to stimulate the interest of visiting professors from several countries, and especially from the USA, to teach in Graz for some time, under exchange programs or some other projects. Administrative approval was not very easy to get at times, because of the increase in teaching staff at our own department. Among the visitors have been Mike Harnish (Tucson), George Kerner (Michigan), Christoph Nyiri (Budapest), Barry Smith (Sheffield), Risto Hilpinen (Turku), and Gershon Weiler, Zvie Bar-On, and Joseph Agassi from Israel. Our department's teaching staff was to a great extent recruited from its own students. In Austria, as in Germany, one can qualify as a university lecturer by the recommendation of a habilitation supervisor, so I have been approached rather often by several candidates. Thus, in the course of 25 years, ten former students of mine (including Christiane Weinberger, Heiner Rutte, Werner Sauer, and Wolfgang Gombocz) obtained the right to lecture here.

With a symposium on Alexius Meinong in 1970, I started a series of international congresses opening Graz to the philosophical world—bordering on the states of the Eastern Block, it had been isolated for a long time. The list of participants included Findlay, Chisholm, Hintikka, Lambert, Marc-Wogau, Küng, and Poser; altogether there were involved philosophers from six countries, with Ryle delivering the opening speech. However, one of my main concerns was the edition of Meinong's collected works, which I later carried out together with Kindinger and Chisholm. In the first ten years of my Graz period, I organized four more symposia, one can almost say five, if one adds a symposium in Göttingen, Germany, in 1977, dedicated to Leonard Nelson[5], with a very large number of participants from Graz, like for instance Chisholm, Haller, Körner, Lehrer, Sauer, Weinberger, and Weinke.

Especially worth mentioning is the Schlick-Neurath symposium in Vienna in 1982 at the Wittgenstein house, an event that gave rise to the re-evaluation of Neurath's philosophy. We were proud of having here C. G. Hempel and Tscha Hung as two foreign members of the Vienna Circle, and a younger generation of philosophers with F. Barone, D. Davidson, F. M. Black, R. M. Chisholm, R. Hilpinen, W. Künne, H. Lauener, K. Lehrer, P. Neurath and J. Vuillemin, who introduced a new analytic spirit into the interpretation of the Vienna Circle philosophy. I will not make a long list of all philosophical undertakings, but only mention the last one in a series of symposia, namely the Meinong Conference of 1995. Its contributions are contained in Volume 50 of *Grazer Philosophische Studien*, the most comprehensive one of all GPS volumes.

There only remains to refer to one topic, which most of my historical research work has been dedicated to: the rediscovery of Austrian philosophy, whose existence had frequently been concealed or denied. Just as not everything written in German can be regarded as German literature, so one cannot simply assume that philosophizing in the same language

5. P. Schroeder, ed., *Vernunft, Erkenntnis, Sittlichkeit. Intern. Philos. Symposium Goettingen, 27.–29. Okt. 1977, aus Anlass des 50. Todestages von Leonard Nelson*. Hamburg 1979.

means to belong to the same movement. Austrian philosophy developed during the Austrian empire about a hundred years ago, its main subjects being of a strictly scientific and anti-Kantian nature with a strong focus on the critique of language. This was a distinctive movement that gradually distanced itself from the more traditional philosophy that prevailed in other German speaking countries.

In 1982, I established the Austrian Philosophy Documentation Centre, whose researchers collect and compile the literary remains of Austrian philosophers, including Brentano's private library and several of his personal documents as well as materials concerning the Vienna Circle and its early history. The research work done by the members of the Documentation Centre resulted for instance in R. Fabian's edition of Christian von Ehrenfels's works, the projected edition of the correspondence of Moritz Schlick, and the works of Oskar Kraus and Alois Höfler.

There is one thing in philosophy that I consider to be of particular importance: not to disregard the role played by the ways or forms of life in which we are rooted, when we ask why our justificational attempts need to come to a halt. Mutual understanding is a basis for dissent as well as for agreement. It may be found in our common sense as well as in our common ways of acting. Without such a basis, or at least the hope for reaching it, we cannot rely on the assumption that what we say, like what we do without saying, will be really understood (and accepted) by others.

If I were forced to summarize my work in philosophy, I would stress two themes. In philosophy proper I tried to find an appropriate answer to questions concerning the kinds and rationality of scientific progress, continuing some of the neglected Neurathian lines of thinking. It seemed to me that theories have more in common with fictional objects than people normally think. In aesthetics I found the question of novelty versus originality an interesting one, especially with regard to new kinds of what we call art. In epistemology a reconstruction of the common-sense experience seemed to be fruitful, also for an explanation of the objects of scientific theories. For some time I have been working on a much larger systematic project with the key-notions *language-mind-world*. Whether I shall ever finish it, remains an open question.

The other theme in my work were studies in the history of philosophy and interpretations of other philosophers' works. Here most significant to me was, firstly, the rediscovery of Austrian philosophy and all my work concerned with it, secondly, the studies on the philosophy of the Vienna

Circle and the proposal of a completely new way to see and to evaluate it, especially the work of Neurath, and, thirdly, my studies on the philosophy of Wittgenstein.

GPS

Subscription Rates
There is no annual subscription rate. Each volume has its own price, varying from 60 to 90 US Dollars. Individual subscribers get a 50% discount.

Information for Contributors

GPS publishes articles on philosophical problems in every area, especially articles related to the analytic tradition. Each year at least two volumes are published, some of them as special issues with invited papers. Reviews are accepted only by invitation.

Manuscripts in German or English should be submitted electronically as an e-mail attachment, either in MS Word or in rtf format, prepared for anonymous reviewing (i.e. without the author's name and affiliation), together with an English abstract of 60–100 words. Footnotes should be kept to a minimum, and references should be incorporated into the text in (author, date, page) form. An alphabetical list of references should follow the text.

A submitted paper will normally be sent to a referee.

Authors are responsible for correcting proofs. Corrections that deviate from the text of the accepted manuscript can be tolerated only in exceptional cases. Authors will receive a free electronic offprint of their article.

Manuscripts should be sent to the following address:
martina.fuerst@uni-graz.at

Email Addresses
Editors
johannes.brandl@sbg.ac.at
marian.david@uni-graz.at
maria.reicher-marek@rwth-aachen.de
stubenberg.1@nd.edu
Managing Editor
martina.fuerst@uni-graz.at

GPS online
Further information about GPS (back volumes, electronic publication, etc.) is available at the publisher's web site: http://www.rodopi.nl, section "Series & Journals".