

# Varieties of Logical Form

**Mark Sainsbury**  
University of Texas at Austin

DOI: 10.2478/disp-2020-0011

BIBLID [0873-626X (2020) 58; pp.223–250]

## **Abstract**

The paper reviews some conceptions of logical form in the light of Andrea Iacona's book *Logical Form*. I distinguish the following: logical form as schematization of natural language, provided by, for example, Aristotle's syllogistic; the relevance to logical form of formal languages like those used by Frege and Russell to express and prove mathematical theorems; Russell's mid-period conception of logical form as the structural cement binding propositions; the conceptions of logical form discussed by Iacona; and logical form regarded as an empirical hypothesis about the psychology of language processing, as in the Discourse Representation Theory tradition. Whereas neither schematization, nor the use of special languages for mathematics, raise general methodological or empirical difficulties, other conceptions of logical form raise at least apparent problems.

## **Keywords**

DRT, Frege, Iacona, logical form, Russell.

logical forms are not edible or hostile, and attention to them is not a cause of longevity. This sufficiently explains why it is only a few eccentric persons, unusually relieved from the struggle for existence, whose attention wanders to such unimportant objects.

Russell (1913: 133)

## 1 Schematization

The origins of the notion of logical form trace back to the realization, clearly present in Aristotle and no doubt earlier, that the value of some reasoning lies in the *pattern* it instantiates, independently of

some of its specific features. Aristotle's syllogistic developed the idea in a sophisticated way. To specify a pattern, you start with sentences of natural language and systematically replace some expressions with schematic letters. Shifting from Aristotle's Greek to English: "All men are mortal" might become "All As are B". We can call the results of schematization "logical forms". They are schematized versions of natural language sentences. Even in such a simple case, some tolerance is required. The schema uses "As", which does not strictly schematize "men" (which contains no "s"), so even at the most basic level we need to recognize, and abstract from, such tiresome and logically irrelevant matters as the different ways in which plurals can be formed.

A series of these logical forms can represent the form of an argument. Interesting series of this kind represent valid "argument forms": logical forms that are valid in the sense that every uniform replacement of schematic elements by appropriate words from the natural language is a valid argument in the natural language. Such arguments can be called valid in virtue of their logical form.

This approach to providing a systematic account of good or valid reasoning is not fully satisfactory, for familiar reasons. One is that there are arguments that seem to be valid in virtue of their form but do not fit any of the patterns made available by the Aristotelian syllogistic. A familiar example:

Every horse is an animal.  
So every horse's head is an animal's head.

This does not directly<sup>1</sup> fit any of the forms recognized in Aristotle's syllogistic, which therefore, at a minimum, needs to be extended. A further dimension of incompleteness, addressed by the Stoics, was the syllogistic's inability to formalize arguments valid in virtue of their sentential form, like *modus ponens*.

Another reason for dissatisfaction is that some supposedly syllogistically valid argument forms have invalid instances. An example is:

Every A is a B. Every B is C. So every A is C.

Geach (1972: 493), crediting his medieval predecessors, showed its

<sup>1</sup> According to Read (2012), Buridan's system proposed accounting for the validity by treating the argument as enthymematic.

invalidity by putting “number” for “A”, “number or its successor” for “B” and “even” for “C”.

Moreover, many have thought that it was a mistake to treat universal quantification in natural language as existentially committing.

Although schematizing may not be a fully effective way of systematizing logical relations among sentences in natural language, it is methodologically straightforward, and brings to light two features that survive into many more recent discussions. One is that it determines a set of logical constants: expressions not to be schematized. The other is that it is a clearly relational notion: “X schematizes Y” expresses a relation that holds between some schematized version of a natural language sentence and the sentence itself. In recent discussions, “X is the logical form of Y” is often thought of as a relation between a sentence in an artificial “formal” language, and a sentence of natural language. This relation is much more problematic.

## 2 Perfect languages

Schematizing is the essence of the syllogistic, so that logical form is thought of as a relation between a natural language sentence and something else. In the late nineteenth century, a very different approach to logic began to emerge. Mathematicians like Venn, Boole, Peano, Dedekind and Schröder began to search for entirely new means of expressing mathematical and logical facts. As Lasswitz (1879/1972: 210) puts it:

It is a long-known complaint that ordinary language is a most inadequate expression of thought; and thus the wish (which already is evident in Leibniz’s attempt at a pasigraphy) that thinking be freed from these chains of language asserts itself again and again.

Frege’s *Begriffsschrift* is a ground-breaking attempt to attain this freedom from natural language, and it has left a permanent mark on logic. The Fregean perspective does not schematize, and it turns its back on natural language, rather than trying to describe it. If there is a notion of logical form to be had within a Fregean perspective, it is not a relation to sentences of natural language. If there is a relation, it is between formulae of the *Begriffsschrift* and thoughts, regarded as abstract objects.

Frege's attitude to natural language emerges in his introduction to the *Begriffsschrift*:

I believe I can make the relation of my "conceptual notation" to ordinary language [*Sprache des Lebens*] clearest if I compare it to the relation of the microscope to the eye. (1879/1972: 104–5)

His idea is that the eye is a superior instrument for some purposes, the microscope for others. Both are trained on reality, and are good for accessing different aspects of it. If Frege thought that the main role of his *Begriffsschrift* was to give the logical form of sentences in natural language, the analogy would require that the microscope's main role is to give information about the human visual system, which is absurd. Rather, there is a common reality to which both instruments are responding, some aspects of which are best accessed in one way, some in another. The aspect of reality that concerned Frege was mathematics, and Frege is claiming that thought on this topic is much better accessed using his *Begriffsschrift* than by using natural language.

The idea emerges clearly when applied to the subject–predicate distinction:

logic up to now has always confined itself too closely to language and grammar. In particular, I believe that the replacement of the concepts of *subject* and *predicate* by *argument* and *function* will prove itself in the long run. (1879: 107)

A distinction between *subject* and *predicate* does *not occur* in my way of representing a judgment.... In this I strictly follow the example of the formula language of mathematics, in which, also, one can distinguish subject and predicate only by doing violence [to the language]. ... We see that here we cannot speak of subject and predicate in the usual sense. ... In my first draft of a formula language, I was misled by the example of [ordinary] language into forming judgments by combining subject and predicate. I soon became convinced, however, that this was an obstacle to my special goal ... (1879: 113)

The *Begriffsschrift* aims to free mathematical thought from the misleading accretions with which it is burdened by its expression in natural languages (1879: 106). Frege is not attempting to use his new notation to clarify or model sentences of natural language; rather, the

plan is to replace natural language by a medium of thought that is superior for mathematical purposes.<sup>2</sup> We may regard the Begriffsschrift as supplying logical forms, but if they are forms of anything, they are forms of thought, not of language. Moreover, unlike the schemata of the syllogistic, Begriffsschrift formulas are not schemata, but are capable of being used to make assertions (judgments). One could not expect otherwise for a language designed to express mathematics. So we have an approach that is quite different from the schematizing approach of the syllogistic, and also different from current proposals to use logical forms to give insight into natural language.

For Russell and Whitehead in *Principia Mathematica* (1910–13), the aim is likewise to provide a language for mathematics, one they hoped would enable the subject to be axiomatized. Hence the formulae are statements, true or false, rather than schemata. Seemingly free variables in theorems of their logic are to be regarded as implicitly “universally closed”: bound by widest-scope universal quantification. They introduce the idea informally:

when a text-book of logic asserts that “ $A$  is  $A$ ”, without any indication as to what  $A$  may be, what is meant is that *any* statement of the form “ $A$  is  $A$ ” is true. (1910–13: 4)

Applying the same idea to their own second “primitive proposition” (“ $p$  or  $p$  implies  $p$ ” (\*1.2)), they say:

Here  $p$  may be *any* elementary proposition: by leaving  $p$  undetermined,

<sup>2</sup> Iacona speaks of Frege’s “hypothesis that sentences [of natural language] have function-argument structure” (2018: 12). He writes that, according to Frege, a sentence of natural language like “Aristotle is rich” has the function-argument structure expressed by “*Rich*(Aristotle)”. This interpretation is based (2018: 12n3) on a later work (“Function and Concept”), in which Frege says that the concept *function* can be “extended” (1891/1984: 147) to include parts of sentences which can be regarded as unsaturated, and so can be extended to include predicates. In this paper, Frege’s remarks on thus extending the notion of function occupy only a single paragraph.

It’s certainly true that Frege’s symbolism has been used by subsequent authors to throw light on natural language. Iacona sometimes somewhat misleadingly attributes this interest to Frege himself, whereas, in my view, Frege’s target was thought. Iacona’s claim “Frege suggested that there is a crucial difference between (1) [‘All philosophers are rich’] and (2) [‘Aristotle is rich’]” (2018: 100) is best interpreted as true of subsequent readers of Frege, rather than of Frege himself.

we obtain an assertion which can be applied to any particular elementary proposition. Such assertions are like the particular enunciations in Euclid ... (1910–13: 93).

Schematization would defeat the aim of stating arithmetic facts. As in the case of Frege's *Begriffsschrift*, the formulae of *Principia* are not "forms" in the sense of something that can be realized by different "matter": different fact-stating sentences. They are already fact-stating sentences.

Like Frege, Russell's primary interest as a logicist is to do better than ordinary language, not to describe it.<sup>3</sup> On the other hand, he and Frege are frequently thought to be the progenitors of our current notions of logical form, nowadays regarded by many as schemata to be used in characterizing ordinary language, and thus having features and aims that Frege and Russell's formal languages lacked. How did this radical shift occur?

### 3 Logical form in early Russell

It's natural to regard the main claim of Russell's "On denoting" (1905) as a logical form proposal: the logical form of a definite description is that of a quantifier phrase, not a referring expression. No doubt that is what many undergraduates are taught nowadays. But, given that the phrase "logical form" does not occur in that paper, and was (so far as I know) not in use by Russell at that time, this interpretation cannot be taken for granted. A more modest idea is that Russell is giving a general account of the truth conditions of sentences containing definite descriptions, and this is a project that need not say anything about logical form. For example, a theorist who makes the claim "'He did it of his own free will' is true iff he could have done otherwise" is making a claim about truth conditions, but may not be making a claim about logical form. One can specify a sentence's truth conditions using a sentence of a distinct logical form (trivial example: the logical form "p" differs from the logical form "p & p", but there is no difference in truth conditions).

Iacona writes that Russell's theory, as presented in "On denoting",

<sup>3</sup> "[We ought not], in our attempts at serious thinking, to be content with ordinary language, with its ambiguities and its abominable syntax" (1951: 694).

shows that “the logical properties of a sentence that contains a definite description ... can be elucidated if the sentence is paraphrased in the appropriate way” (2018: 18). Russell does not use the words “paraphrase”, but he does speak of meaning, abbreviation, and also of the “reduction of all propositions in which denoting phrases occur to forms in which no such phrases occur” (1905: 482). Perhaps reduction is what Russell has in mind when he speaks of “becoming” in the following:

Thus “the father of Charles II was executed” becomes:—

A: “It is not always false of  $x$  that  $x$  begat Charles II and that  $x$  was executed and that ‘if  $y$  begat Charles II,  $y$  is identical with  $x$ ’ is always true of  $y$ .” (1905: 482)

Should we regard A, and similar reductions, as merely trying to make the truth conditions explicit, or as, in addition, trying to make the logical form fully explicit? The answer has to be the former, for Russell gives specifications of truth conditions which are so different that they cannot all be regarded as giving a fully explicit presentation of the logical form.

“Scott was the author of *Waverley*” is treated as equivalent to each of the following:

B: One and only one entity wrote *Waverley*, and Scott was identical with that one. (1905: 488)

C: It is not always false of  $x$  that  $x$  wrote *Waverley*, that it is always true of  $y$  that if  $y$  wrote *Waverley*  $y$  is identical with  $x$ , and that Scott is identical with  $x$ . (488)

The latter is said to be “wholly explicit”. A fourth version is suggested by the discussion in *Introduction to Mathematical Philosophy* of “The author of *Waverley* was Scotch”:

D: (1) at least one person wrote *Waverley*  
 (2) at most one person wrote *Waverley*  
 (3) whoever wrote *Waverley* was Scotch. (1919b: 331)

Concerning D, he concludes that “the three [quoted propositions] may be taken as defining what is meant by the proposition ‘the au-

thor of *Waverley* was Scotch” (1919b: 331).

The various formulations of the theory of descriptions show that the use of logical symbols is not needed for the expression of the theory, which some might regard as enough to preclude them from being logical forms. Admittedly he sometimes uses what might be thought of as variables, but these are really mere schemata, a way of attaining generality:

speaking generally, suppose we wish to say that the author of *Waverley* had the property  $\phi$ , what we wish to say is equivalent to “One and only one entity wrote *Waverley*, and that one had the property  $\phi$ ”. (1905: 488)

These words give no reason to suppose that anything more is happening than a specification of truth conditions: *equivalence* is the central relation, and this is symmetric, whereas *being the logical form of* is not.<sup>4</sup>

Russell’s four versions A–D are so different from one another that they cannot all be regarded as making a unique logical form fully explicit. A is a single sentence, whereas D consists in three sentences (propositions) not even conjoined by “and”. The first three explicitly invoke identity, the fourth does not. Unlike the other versions, A involves internal quotation; and some but not all versions use the notion of “not always false”. On some views (reasonable, but in my view incorrect), all four glosses count as having the same logical form. But they cannot all be regarded as *explicitly specifying* a single logical form. The relation *x has the same logical form as y* is an identity relation and so reflexive and symmetric. Let’s for a moment roll with the assumption that it holds of every pair in the list A–D. By contrast, *x is the logical form of y* is not symmetric. If logical forms are unique, no two sentences (or formulas) are the logical form of any one sentence.<sup>5</sup> If Russell was

---

<sup>4</sup> Elsewhere in the paper, Russell quantifies over forms, for example “a phrase is denoting solely in virtue of its *form*” (1905: 479), and denoting phrases are the topic of the paper. This is intended as merely a grammatical generalization, preempting the thought that every “denoting phrase” denotes. It does not mark a notion deserving the phrase “logical form”.

<sup>5</sup> They are not strictly speaking unique, for every sentence has the form “p” (so there are various “depths” of logical analysis), and merely alphabetical variation among formulae is discounted. Even so, many theorists automatically preface “logical form” with “the”. Iacona discusses some related issues around p.



trying to specify logical forms, he would not have missed this point. Hence we have to conclude that in these passages this was not his intention. But he was certainly attempting to identify truth conditions.

For those who regard the 1905 account as a logical form proposal, I suggest the following exercise. Consider the claim that the contribution to truth conditions made by a definite description is that of a uniqueness quantifier. In more Russellian language, the claim is that sentences of the form “The F is G” are true iff there is exactly one F and it is G.<sup>6</sup> Since this is merely a statement of truth conditions, available to someone who thought the notion of logical form too confused to be usable, those who regard Russell’s view as a logical form proposal must regard it as leaving something out. So: make explicit what is left out, and then point to the sentences in Russell’s article in which he affirms this omitted material. Nothing less would constitute a full vindication of the view that “On denoting” is intended to specify not just truth conditions but, in addition, logical form.

The theory of descriptions of 1905 is not the place to look for Russell’s conception of logical form. An early explicit use of the phrase “logical form” occurs in his account of ordinary proper names in *Principia Mathematica*:

A proposition such as “Apollo exists” is really of the same logical form [as “the most perfect Being exists”], although it does not explicitly contain the word *the*. For “Apollo” means really “the object having such-and-such properties . . . .” (Russell and Whitehead 1910–13: 31)

Since the example of “Apollo” is used to make the same point in “On denoting”, it provides some evidence that it’s appropriate to treat the early theory as a logical form proposal. If so, we might feel at a loss to explain what logical form amounts to for the early Russell. “Apollo” does not explicitly contain “the”. But what sense can we make of the

64. There are questions about what to make of logical equivalences (would it be wrong to formalize “all” as “ $\neg\exists\neg$ ”?), and questions about whether the search for the “deepest” or “most revealing” formalization should, for example, represent “drake” as conjunctive, made up of “male” and “duck”.

<sup>6</sup> For simplicity, I omit the different claim required for sentences of the form “The F exists”.

idea that it does so implicitly?<sup>7</sup> If we go forward barely a year, to 1911, we find an explanation:

Common words, even proper names, are usually really descriptions. That is to say, the thought in the mind of a person using a proper name correctly can generally only be expressed explicitly if we replace the proper name by a description. Moreover, the description required to express the thought will vary for different people, or for the same person at different times. (1910-11: 114)

If it is legitimate to treat the two quotations just displayed as belonging to a unified stream of thought (admittedly, a dubious assumption for someone as mercurial as Russell), we could regard his theory of ordinary proper names as a logical form proposal. The proposal makes logical form an asymmetric relation, so it is closer to current conceptions than mere equivalence of truth conditions. It raises the question of whether or not the various thoughts mentioned in the second quotation, differing from person to person, have the same logical form. If one says they do not have the same logical form, then logical form differences are created just by differences in individual psychology. If one says that they do have the same logical form, one appeals to a conception of logical form differing both from Iacona's and from currently standard views.<sup>8</sup>

At this stage in Russell's development, he is not operating with a thought-out and coherent notion of logical form. If we attribute logical form doctrines to his work before 1912, we risk anachronistic interpretation.

<sup>7</sup> As Iacona says: "there is an important sense in which logical form may be not visible in surface structure" (2018: 35); perhaps "visible" should be prefaced by "readily".

<sup>8</sup> The different definite descriptions in the minds of different users of a name may themselves differ in logical structure. Suppose speaker A associates "de Gaulle" with "The President of France in 1959" and speaker B associates it with "The man with a nose longer than any other in Europe in 1959". A's utterance of "De Gaulle is F" is animated by the thought that the President of France in 1959 is F. The utterance will prompt B to think that the man with a nose longer than any in Europe is F. Many theorists, including I think Iacona, will hold that B's thought has a logical form that A's thought lacks: e.g. B's has universal quantification over noses, which A's lacks, so the number of quantifiers in B's thought exceeds the number of quantifiers in A's thought; hence they differ in logical form.

#### 4 Middle period Russell on logical form

In his 1914 lecture entitled “Logic as the essence of philosophy”, Russell describes logical form as a relational notion, something possessed by sentences of natural language. Those who understand natural languages need to appreciate this property:

In every proposition and in every inference there is, besides the particular subject-matter concerned, a certain form, a way in which the constituents of the proposition or inference are put together. ... It is obvious that the knowledge of logical forms is something quite different from knowledge of existing things. ... Thus some kind of knowledge of logical forms, though with most people it is not explicit, is involved in all understanding of discourse. It is the business of philosophical logic to extract this knowledge from its concrete integuments, and to render it explicit and pure. (1914: 53)

Logical form is the glue that binds the non-logical constituents of propositions together. A specification of logical form displays the glue without specifying the parts that are glued. This is closer to what might be called the modern notion of logical form. Logical form needs to be *known* in understanding natural language, but the relevant knowledge is typically *implicit*. Philosophical logic makes it *explicit*, by representing logical form in a distinct language, one having the requisite explicitness and purity. The language of *Principia Mathematica* was ready for this new deployment: its formulae are no longer specifically in the service of mathematics, but are reconfigured as schemata that can help explain features of natural language: the requirement of universal closure is quietly dropped. This marks a profound shift.

At close to the time of the lecture from which I just quoted, Russell had been working on the *Theory of Knowledge* manuscript, not published in his lifetime (generally dated to 1913). He quite casually introduces the phrase “logical form” early on, claiming:

It can be shown that a judgment, and generally all thought whose expression involves propositions, must be a fact of a different logical form from any of the series: subject-predicate facts, dual relations, triple relations, etc. In this way, a difficult and interesting problem of pure logic arises, namely the problem of enlarging the inventory of logical forms so as to include forms appropriate to the facts of epistemology. (1913: 46)

A variety of logical forms are mentioned, but epistemology introduces a new beast for our zoo (as he put it later: 1919a: 60), since it involves attributions of belief and knowledge: what are now often called propositional attitude attributions.<sup>9</sup>

Chapter 9 of the *Theory of Knowledge* manuscript is devoted to logic, and its initial account of logical form, as glue, would cause no surprises in a text written today. Logical form

is the way in which the constituents are combined in the complex [proposition]. It is such pure “forms” that occur in logic. ... When we have reached the form “ $xRy$ ”, we have effected the utmost generalization which is possible starting from “Socrates precedes Plato”. (1913: 98)

Here a form is specified in familiar notation, and it is plain that different sentences may share a form. Explaining natural language and thought requires identifying forms in a way that is now familiar.

Here there clearly emerges, perhaps for the first time, a conception of logical form that really could count as a “property of sentences” (Iacona: 2018: 1), as opposed to a property of facts or thoughts.<sup>10</sup> But it comes with some specifically Russellian baggage, notably the view that logical forms are the objects of acquaintance:

It is difficult to see how we could possibly understand how Socrates and Plato and “precedes” are to be combined [in the proposition “Socrates precedes Plato”] unless we had acquaintance with the form of the complex. As a matter of introspection, it may often be hard to detect such acquaintance; but there is no doubt that, especially where very abstract matters are concerned, we often have an acquaintance which we find it difficult to isolate or to become acquainted with. (1913: 99)

Russell’s position at the time of World War I has four features in common with a modern conception of logical form: it involves a classification of forms on familiar lines (“subject-predicate”, “relation”, “dual complex”), it uses formulae like  $xRy$  as schematic (and so in a way quite unlike their use in *Principia*), it singles out the logical con-

<sup>9</sup> Given that the zoo was already quite well populated, it is a little misleading to say, as Iacona does (2018: 21): “Russell’s use of the term ‘logical form’ is motivated by considerations that pertain to his multiple-relation theory of judgment”.

<sup>10</sup> We might wonder whether Russell’s earlier remark (1904: 59) that logic is not concerned with words but with what they stand for holds for his 1913 conception.

starts in natural language (words like “or”, “not”, “all” and “some” (1919a: 101)), and it regards logical form as sometimes only inexplicitly present in natural language sentences, thus giving rise to what might nowadays be called a “research project”: find the logical forms.

## 5 Two conceptions of logical form?

In his crisp and elegant book *Logical Form*, Iacona’s main claim is that the uses to which the notion of logical form is put nowadays require two distinct notions:

the syntactic notion suits the semantic role, although it is unsuitable for the logical role ... the truth-conditional notion suits the logical role, although it is unsuitable for the semantic role. (2018: 68)

The different roles are introduced as follows:

the *logical role* “concerns the formal explanation of logical properties and logical relations” (2018: 37).

The *semantic role* “concerns the formulation of a compositional theory of meaning” (2018: 37).

This is initially a surprising contrast. One would expect a syntactic notion to suit the logical role, for logic, with its logical constants, formulas and proofs, is essentially a matter of syntax: everything that matters to the extension of the single turnstile. And one would expect a truth-conditional notion to suit the semantic role, since truth conditions are at the heart of semantics. A further possible source of surprise is that the semantic role strongly suggests Davidson’s program, but it’s hard to see any demand, in Davidson’s account of logical form, that logical forms and what they are forms of should be syntactically similar, so it would be surprising if Davidson’s program required a syntactical notion of logical form. Indeed, when one thinks of Davidson’s account of action sentences, which Iacona discusses at p. 36, one could hardly suppose that Davidson (or anyone else) would think that the proposed conjunctive “logical form” matches the syntax of the natural language original.

In understanding Iacona’s perspective, I found it helpful to consider the importance he attaches to context-sensitive expressions, like pronouns. On his view, the *meaning* of a given pronoun is invariant,

but its *reference* is not, and it is thanks to its reference that it contributes to truth conditions. Then an account of meaning will fail to do justice to reference (in the case of pronouns), and so will fail to do justice to truth conditions. This is idiosyncratic, for Davidson certainly thought that a compositional theory of meaning had to do justice to reference, though he recognized, starting in 1967 with the famous remark about the “fly in the ointment” (318), that it was not obvious how best to incorporate contextually sensitive reference into his account.

Davidson argued that a theory of truth for a natural language would not merely specify its compositional semantics, but would also settle which of its arguments are logically valid: they would be those object language arguments for which a theorem of the metalanguage says that if its premises are true, so is its conclusion. On this view, a correct semantic theory for a language would deliver a correct account of logical validity in that language, and so could not afford to neglect logical form, however understood. For Davidson, the very same truth-theoretic procedure would deliver both truth conditions and *formal* validity. Given Davidson’s attempts to incorporate context-sensitivity, it would be too quick to dismiss this unifying conception of logical form merely by claiming that the attempts failed. A detailed discussion would be required.

Logical form is a relational notion: X is the logical form of Y. There is a wide choice of first term: Frege’s *Begriffsschrift*, Russell and Whitehead’s *Principia Mathematica*, the language of Zermelo-Frankel set theory, sentential calculus, predicate calculus with or without identity and with only binary or only unary quantifiers or both, the Polish notation favored by Geach, Quantified Modal Logic, alethic logic, and so on. For most of the book, Iacona offers classical first order forms, treated as schemata, though he also uses English annotated with variables for traces, which he regards as exemplifying LF (these are non-schematic), a candidate source of logical forms. (The last two chapters appeal to set theoretic symbols with symbols for cardinality-related notions.) There is also a choice of logics in the sense of axioms and rules: classical logic, the logic of *Principia Mathematica*, relevance logic, paraconsistent logic, free logic, and so on. Iacona chooses classical logic, and, like the various symbolisms,

alternatives are irrelevant to his project.<sup>11</sup>

Iacona suggests that the second term of the logical form relation is determined by *content*, where content is what is said (or would be said) by the utterance of a sentence on a specific occasion. “Determined by” ensures that difference of content must be reflected by difference of logical form. Content is an occasion-sensitive notion: utterances of a sentence like “I am hungry” in different contexts have different contents, and so call for different logical forms. This is one of Iacona’s most original contributions.

Traditionally, context is simply ignored in discussions of logic. A sentence like “That is red” would be awarded subject-predicate logical form, using for example the formula “*Fa*”, disregarding the fact that the English sentence may be used to say different things in different contexts, and so has context-sensitive content. The standard approach likewise regards all sentences like “Plato is a philosopher”, “Aristotle is a philosopher”, “John is happy”, “Mary is tall” as sharing the subject-predicate form, and so all as equally well formalized as *Fa*. Iacona points out that if we are formalizing a *series* of sentences, perhaps because we are testing an argument for validity, this free-wheeling approach needs to be constrained: if an expression occurs more than once in an argument, it must be formalized by the same formal expression. Otherwise validity-relevant features may be overlooked. If the first question of the logic test says “Formalize ‘Aristotle is a philosopher’”, “*Fa*” is a correct answer, as it would also have been to the question “Formalize ‘Napoleon was an emperor’”. “*Gb*” would also have been correct in either case. But if the question demands the student to formalize the argument “Aristotle is a philosopher. Napoleon is an emperor. So there is at least one philosopher and at least one emperor”, different individual constants must be used for “Aristotle” and “Napoleon”, and different predicate letters for “philosopher” and “emperor”. We can only evaluate logical form proposals relative to a set of sentences and a correspondence scheme, explaining which words of the natural language correspond to which formal expressions.

At the start of “On sense and reference” Frege says:

<sup>11</sup> It would be interesting to know Iacona’s views about what makes for the right logic for logical forms.

$a=a$  and  $a=b$  are obviously statements of differing cognitive value;  $a=a$  holds a priori and, according to Kant, is to be labelled analytic, while statements of the form  $a=b$  often contain very valuable extensions of our knowledge and cannot always be established a priori. (1892: 157)

What he says about “ $a=b$ ” presupposes that distinct names may share an object. What he says about “ $a=a$ ” presupposes that distinct objects cannot share a name. Thus baldly stated, thinking of the individuation of names in a typographical way, the first presupposition is true but the second is false. Yet name-sharing by objects, and object-sharing by names, seem to be the same kind of phenomenon: it would be hard to imagine a system of signs with one feature but not the other, unless one was prohibited by some explicit convention. Nowadays, Frege’s discrepant treatment is generally taken for granted in logical form proposals. Is it justifiable to do so?

An important feature of Iacona’s book is that he challenges this orthodoxy (§6.3, §6.4), led there by his conception of logical form as a property of content, including context-dependent content. He proposes two rules (2018: 71, 74).

(N1) in logical form, distinct objects must be denoted by distinct names

(N2) distinct names must denote distinct objects.

This makes his position symmetric in a way that Frege’s is not: in logical form, both object-sharing by names and name-sharing by objects are prohibited. The consequence is that the process of formalization is not apriori. One cannot know apriori whether “Hesperus is Phosphorus” should be formalized as “ $a=a$ ” or as “ $a=b$ ”. The first is the unique correct formalization given that Hesperus is Phosphorus, otherwise one would infringe (N2); and this cannot be known apriori.

Iacona bases this position on its explanatory power (2018: 71). It connects with a claim about the nature of knowledge of logical form, and he suggests that the view does justice both to some questions about the logical form of context-sensitive expressions, and to some familiar puzzles about belief.

He envisages an utterance of “This is different from this” in which the different occurrences of “this” refer to different things. Logicians



have largely neglected the issue of how such sentences should be formalized. A conventional opinion would be that here we run into features of language use that are not within the province of logic. By contrast, Iacona suggests that there is a *formally valid* inference to the conclusion “There are at least two things” (2018: 62), since the two occurrences of “this” must be formalized with different name-letters. Likewise, he suggests that “Plato is different from Plato”, understood as involving different referents for the two name-tokens, should be formalized using different name-letters, so this premise also supports a *formally valid* inference to the conclusion that there are at least two things. The natural contrary view is that the fact about the difference of reference is not a formal one. We should no more think of these inferences as formally valid than an inference from “Donald is a drake” to “Donald is male”. We could pretend this is a matter of formal validity by formalizing the premise as a conjunction, but such a pretense, systematically carried out, would eliminate any distinction between formal and non-formal validity.<sup>12</sup> Some theorists might find this elimination unacceptable; Iacona himself adopts a neutral stance (2018: 66).

Iacona’s position might seem to follow from taking contents (as opposed to sentences) as the possessors of logical form. For example, he says that “knowledge of logical form involves knowledge of reference” (2018: 80). One cannot deny that the content expressed by an utterance of “This is this”, in which different things are referred to, is different from the content expressed by an utterance of the same words in which the same thing is referred to. But for purposes of validity, what one needs to know is not the content itself, which would involve knowing to which object the occurrences of “this” referred, but the structure of the content: whether what is referred to by the two tokens is the same or different. Knowledge of reference is neither necessary nor sufficient for knowledge of sameness and difference of reference. For example, one might know that the first “this” refers to Hesperus and the second to Phosphorus, so one knows the

<sup>12</sup> The “drake” example is from Iacona (65). He says that the fact that a conjunctive formalization of the premise involves conceptual analysis is not enough to show it is not a matter of logical form, since e.g. Russell’s theory of descriptions also involves conceptual analysis. He does not elaborate.

reference of each, and yet fail to know that the occurrences refer to the same thing. And one might know that the two occurrences refer to the same thing without knowing which thing that is. Hence Iacona's two rules do not follow directly from his assumption that content is what logical form is of.

The two rules restore a kind of symmetry that Frege's position lacks, but Fregean asymmetry is easy to justify, and Iacona's symmetry faces some problems.

Stipulating that, in logic, we should not use a name that refers to distinct objects is simply part of the standard requirement that logical notation not be ambiguous. It is easily realized syntactically, for example by taking " $a=a$ " as a theorem schema, and by making the semantic interpretation function one-one (i.e. genuinely a function). To implement the other plan one would have to stipulate that distinct name-letters are never assigned the same object in an interpretation, and take " $a \neq b$ " as an axiom schema. Such an approach has been considered, as Iacona notes, but one might prefer a notion of logical form that did not commit to such an unfamiliar and little-explored system.

Problems facing the view that logical forms cannot always be known apriori include:

- (1) Logic will not be an apriori discipline, despite being generally understood to be one. This understanding is more than a merely arbitrary stipulation. It affects how the subject is taught (conventionally, no astronomy course is required to enable a student to formalize "Hesperus is Phosphorus"), and how arguments in it are conducted. Turning it into a partly non-apriori subject would require a very powerful reason.
- (2) As against Frege, Iacona must say that no sentence whose logical form is " $a=b$ " has cognitive value.
- (3) The view would count the following argument as logically valid:  
     Clark Kent writes  
     Superman flies  
     so someone both writes and flies.

- (4) It would represent the ancient astronomers as logically as well as astronomically ignorant.
- (5) Finally Iacona's position submerges a difference that many have felt between various puzzles cases.

To elaborate this last point: Iacona redescribes various familiar examples (e.g. Kripke's Pierre and Peter, Boghossian's Pavarotti) as arising through failure to know logical forms, since these cannot always be known apriori. In the case of Pierre, "since 'Londres' and 'London' refer to the same city, they must be represented by the same individual constant" (2018: 76), a representation that would reveal Pierre's beliefs to be contradictory ( $Fa$  and  $\neg Fa$ ). But since "he is not in a position to know by reflection the logical form of those beliefs" (2018: 76), he is not in a position to appreciate that they are contradictory. Peter's beliefs, that Paderewski has musical talent and that Paderewski lacks musical talent, are likewise contradictory, though Peter is not in a position to know this apriori.

This elides a potential distinction between the two examples: Peter's case has struck some people as more challenging than Pierre's. Suppose Pierre represents one belief in French ("Londres est jolie") and another in English ("London is not pretty"). By conventional standards, these are inconsistent but not contradictory (contradictories are a proper subset of inconsistencies: inconsistencies having the form " $p \ \& \ \neg p$ "). This is what explains the possibility of rational Pierre having inconsistent beliefs: their inconsistency is not revealed by their logical form, and rationality would always deliver knowledge of logical form. By contrast, it's hard not to convict Peter of irrationality, since his beliefs are apparently contradictory. One response has been to attribute to him distinct representations of Paderewski in his language of thought, thus representing his belief states as inconsistent but not contradictory, like Pierre's. Such a response would be ruled out in advance on Iacona's view.

Iacona's account lacks explanatory value. He has to say that the astronomical discovery about Hesperus and Phosphorus led to a logical form discovery: it turned out not merely that Hesperus was Phosphorus, but that what was discovered was something of the logical form  $a=a$ . But the logical form discovery as such does not explain

anything, and indeed obscures the astronomical achievement. No adequate explanation can avoid appealing to astronomy, but once that appeal has been made, no further explanation is required. The claim about logical form, even if accepted as correct, adds nothing.

Something similar applies to the cases of Pierre and Peter. Peter failed to recognize the inconsistency of his beliefs because he didn't realize that Londres and London were the same city, a geographical matter. Peter failed to realize that his Paderewski beliefs were contradictory because he didn't realize that Paderewski the musician and Paderewski the politician were the same person. Any explanation that omitted these facts would be inadequate. But once the facts are in place, nothing else is needed for a fully satisfying explanation. Of course, the explanation leaves unresolved various questions about the semantics of belief attribution, but so does Iacona's claim about logical form. It's not as if one could happily attribute to either subject, prior to the discovery, beliefs in which the Iaconian logical form was used in the content attributed.

The proposal seems unable to accommodate some more nuanced facts about reference and attribution. For example, it apparently entails

- (1) Pierre believes that London is pretty and believes that London is not pretty.

Normally a conjunction of belief attributions relating to the same person at the same time can be merged without loss of truth, suggesting that these attributions entail:

- (2) Pierre believes that London is pretty and that London is not pretty.

And, normally, a name re-used in a sentence can be replaced by a pronoun, yielding:

- (3) Pierre believes that London is pretty and that it is not pretty.

And, normally, such sentences are equivalent to their contractions:

- (4) Pierre believes that London is pretty and not pretty.

I think that (4) is not entailed by the original puzzle. But I found it hard to figure out how Iacona's proposal could avoid regarding it as a

correct. What would be needed is an explanation of why one of the transitions from (1) through (4) fails to be truth-preserving, but I do not see how this could make use of the notion of Iaconian logical form.

## 6 Skepticism about logical form

Schematic and perfect-language conceptions of logical form have no methodological flaws. Syllogistic patterns are really present, and perfectly visible, in natural language. And if mathematicians like Frege and Russell find value in creating new, special and perfect languages for the expression of mathematical truths, there can be no general objection to their procedure, nor any methodological problem. Such approaches make no claim to reveal anything about natural language.

Quine's use of the notion of logical form raises no methodological problems (at least, not those I am about to raise). That's because the relation of paraphrase he says holds between a natural language sentence and its logical form does not require synonymy, and not even sameness of truth conditions, as can be seen from the use of paraphrase to reduce ambiguity.<sup>13</sup> Frege and Quine make use of *logical form* in similar ways: just as what Frege valued in the Begriffsschrift was not what it could tell us about natural language, but how it could facilitate mathematics, so what Quine valued in "regimentation" was not what it could tell us about natural language, but how it could facilitate expressions of ontological commitment. Frege sought a way of expressing mathematical facts precisely, making it easy to check whether an alleged proof was complete. Quine sought a way of making ontological commitment instantly recognizable. For Quine, paraphrase or regimentation is an opportunity to reconsider what one really wishes to commit to; it is not the start of an investigation into the unhelpful vagaries of ordinary language.

Methodological problems arise when this attitude is replaced by one according to which the process of formalization is supposed to reveal or even explain something about the logic or semantics of natural language. In Davidson's vision, logical form is contrasted with something superficial, requiring special skills to uncloak:

<sup>13</sup> "Typically, indeed, the paraphrasing of a sentence S of ordinary language into logical symbols will issue in substantial divergences" (1960: 159).

Logical form was invented to contrast with something else that is held to be apparent but mere: the form we are led to assign to sentences by superficial analogy or traditional grammar. What meets the eye or ear in language has the charm, complexity, convenience, and deceit of other conventions of the market place, but underlying it is the solid currency of a plainer, duller structure, without wit but also without pretence. This true coin, the deep structure, need never feature directly in the transactions of real life. As long as we know how to redeem our paper we can enjoy the benefits of credit. (1967/1980: 137)

This raises the question of what makes the “true coin” the real thing, and what role it plays in understanding and inferring. Natural language sentences are said to have logical forms presented by formulae that are at least superficially very different; logical forms are said to be easily missed or even “invisible” (an adjective Iacona applies both to the “old” (2018: 24) and the “new” (2018: 34) conceptions of logical form). And their identification, despite their invisibility, is supposed to be a precondition for a properly theoretical understanding of the logic and semantics of their targets in natural language. How can something invisible, or something known only to experts, be the gold that ultimately gives value to the paper money of ordinary idiom?

There are two salient hypotheses (perhaps corresponding to Iacona’s two conceptions of logical form): one is that logical forms bring to the fore the logically relevant features of natural language sentences. The other is that they bring to the fore features of the semantic composition of natural language sentences. The hypotheses are not obviously inconsistent. We saw earlier that Davidson thought that logical forms play both roles. I contend that they are not fitted to play either role.

We all know that it is sometimes much easier to follow a train of reasoning in, say, classical logic than in corresponding sentences of natural language. There are at least three aspects to the greater ease: there is schematization, bringing out when an element recurs and eliminating untidy clutter; there is disambiguation, notably of scope; and there is the use of logical constants with defined logical powers, replacing their possibly logically murky analogs in natural language (conditionals are a clear example). But all these undeniable benefits are available under the supposition that the natural language

originals have simply been replaced by truth conditional equivalents in the formal language. There is no need to suppose that the formalizations provide otherwise unavailable insight into the natural language sentences. It could be, as in the tradition of Frege and Quine, that the logical language serves certain purposes better than does the natural language, so natural language is to be replaced by something more suitable, rather than being made an object of investigation.

Similar considerations apply to the claim that logical form reveals semantic features of natural languages sentences. We already understand our sentences, so what can there be to reveal? For Davidson, the answer was an explanation of the compositional features of language, summed up in providing an answer to his insightful question “What are these familiar words doing here?” (1968: 132). But his use of logical form to answer this question falls manifestly short. For example, his only reason for rejecting irreducible adjectives and adverbs is that they do not feature in first order languages. Adjectival modification is equivalent to predicate or sentence conjunction in some cases (when an FG is something both F and G), but certainly not in all (a small person might be a huge help). Similarly, while some sentences containing adverbs can, with some rigmarole, be given truth conditions in terms of predicates, this is not true of all (walking quickly and turning quickly involve different kinds of quickness). The point is not that Davidson’s project hit roadblocks, but that its very conception is questionable. It could be expressed as a dilemma: If logical forms have the same semantic and logical features as what they are supposedly forms of, why don’t natural language sentences count as their own logical forms? If logical forms differ in their logical and semantic features from what they are supposedly logical forms of, how can they throw light on the semantics and logic of their targets?

Perhaps the question can be answered by an illustrative example. Iacona suggests that the validity of the following natural language argument

(9) Plato is different from Aristotle

(10) There are at least two things

can be explained by saying that it has the following form:

(11)  $a \neq b$

---

(12)  $\exists x \exists y x \neq y$  (2018: 72)

But how does the explanation work? If we were not already convinced of the validity of the argument from (9) to (10) we would not accept (11) to (12) as its logical form. The formula seems no more of an *explanation* than had we translated the original into French.

To take another example, might the conjunctive theory of adjectival modification supply an explanation of adjectival detachment (from *x is FG infer x is G*)? The explanatory value is an illusion, for a needed premise to the explanation is that the relevant adjective is the kind of adjective for which inferences like this hold, that is, an intersective adjective. Nothing is added by representing it as conjunction elimination.

Rather than consider responses to criticisms of this sort, let me turn to a view of logical form to which they are irrelevant.<sup>14</sup> On this view, a logical form proposal is an empirical hypothesis about the actual processing methods that people typically unconsciously use when understanding language. In the words of Kamp and Reyle, a logical form proposal is a

description ... conceived as an (idealized) analysis of the process whereby the recipient of an utterance comes to grasp the thoughts that the utterance contains. (1993: 8)

We might dream of opening the heads of interpreters to find out how they actually do it. But, as Kamp and Reyle say, we would not know what to look for. Meanwhile, we can only appeal to inferences to the best explanation: if the logical form proposal is consistent with what we know about how interpreters actually function, and no better explanation is available, then it is to be accepted, at least provisionally. A Discourse Representation instruction like “if you encounter a new indefinite noun phrase, introduce a new discourse referent” is correct just if this is an idealized version of a rule that interpreters actually obey when they encounter a token of the specified kind of expression. We are justified in regarding it as correct if there is no

<sup>14</sup> Iacona alludes to this program (2018: 38), citing Kamp 2015.



better explanation of interpretive behavior.

This makes good sense of the explanatory value of logical form, both as explaining interpretation and as explaining inference. It addresses psychological inferential processes which genuinely occur, even if they are typically not available to introspection. In this respect the proposal is not merely to provide a theoretical description of a practical ability. That could be done by associating each sentence with an interpretation. But that achievement, however remarkable it would be if it respected Tarskian and Davidsonian constraints, can in itself tell us nothing new: we already know the relevant associations, since we already understand our language. What is interesting about the Kamp–Reyle approach, as I understand it, is that it adds a claim about how interpreters in fact come to know the interpretations of the sentences they understand; DRT has something to say about the actual psychological processes.

This provides a model of how logical form proposals could properly be regarded as explanatory of understanding and inferring in natural language: they are psychological hypotheses. I close, however, by pointing out a possible threat to the approach.

The approach takes for granted that interpretive knowledge is theoretical knowledge, the kind that can be presented as a “theory” in the technical sense of a set of sentences closed under an entailment relation. But that assumption is risky. An alternative account is that we acquire language through “deep learning” mechanisms which ground their skills in training and feedback, rather than in the acquisition of theoretical or factual knowledge. Knowledge of language would be implemented in the style of recent artificial game-playing machines, rather than in the style of axiom-and-theorem based older attempts exemplified by Good Old Fashioned AI.

One sign that deep learning affords a better picture of the inner mechanics of language processing is that we are resistant to information (or misinformation) that would undermine what, on the theory, have to be regarded as premises to the inference. No consciously held opinion about how to respond to, say, a new indefinite noun phrase has the slightest chance of affecting our interpretative practice. This suggests that although a theoretical description of the practice might correctly correlate inputs and outputs, it would not correctly describe the kind of processing that leads from the one to the other.

Although I do not know how this issue could be resolved empirically, there are some pointers favoring the deep learning approach. Language processing is normally very rapid and automatic, like systems that are responsible for identifying people or the sex of day old chicks, and controlling bodily movements (notably in sports), systems typically modeled in the deep learning way. It is markedly unlike what Hume referred to as “the laboured deductions of the understanding” (1748: §5.22). Our capacity to learn a language has a significant window, declining rapidly after around age seven, whereas our capacity to learn new facts and theories increases after that time. And whereas the output of some processes that are quick and automatic can be explained by their possessor on the basis of introspection (for example, even if a farmer need not engage in reflection to detect that an animal is sick, she may well be able to explain the evidence if asked), uncovering the facts supposedly appealed to by our language processing system is a very difficult task. Moreover, language understanding is difficult to undermine: even if we are told on impeccable authority that a conversation we are overhearing is not in English, despite using what appear to be English words in standard English arrangements, we will be unable to prevent ourselves interpreting it in the normal way. Knowing these interpretations to be incorrect does not stop them foisting themselves upon us (just as the inequality of the lines in a Müller-Lyer illusion foists itself upon us, despite our unshakeable knowledge of their equality). This suggests we did not reach our interpretive conclusion on the basis of some premises, for these should have been undermined when we came to believe that what we are hearing is not English. Finally, attempts to create artificial systems of language comprehension that adopt the deep learning approach have already overtaken, in terms of practical success, earlier more axiom-driven approaches. They provide no explanations or justification, but, somehow or other, are good at prediction.

Whatever one may think of the correctness of the processing based model of logical form, it has the merit of bringing into sharp focus a notion of logical form that justifies many of the traditional views: it would be genuinely explanatory both of interpretation and inference, would appeal to hidden elements, and would be capable of making surprising predictions.

Mark Sainsbury  
University of Texas at Austin

### References

- Davidson, D. 1967. Truth and meaning. *Synthese* 7: 307–23.
- Davidson, D. 1967/1980. Criticism, comment and defense. Reprinted in his *Essays on Actions and Events*. Oxford: Clarendon Press, 1980: 122–48.
- Davidson, D. 1968. On saying that. *Synthese* 19: 130–46.
- Frege, G. 1879. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle, L. Nebert. Page numbers refer to *Conceptual Notation and Related Articles* translated and edited by T. W. Bynum. Oxford: Oxford University Press, 1972.
- Frege, G. 1891. Function and concept. In *Collected Papers on Mathematics, Logic and Philosophy*. Edited by Brian McGuinness. Oxford: Basil Blackwell, 1984: 137–56.
- Frege, G. 1892. On sense and meaning. In *Collected Papers on Mathematics, Logic and Philosophy*. Edited by Brian McGuinness. Oxford: Basil Blackwell, 1984: 157–77.
- Geach, P. T. 1972. A program for syntax. In *Semantics of Natural Language*. *Synthese Library vol 40*, edited by D. Davidson and G. Harman. Dordrecht: Springer: 483–97.
- Hume, D. 1748. *An Enquiry Concerning Human Understanding*. London: Andrew Millar.
- Iacona, A. 2018. *Logical Form*. Dordrecht: Springer Synthese Library.
- Kamp, H. 2015. Using proper names as intermediaries between labelled entity representations. *Erkenntnis* 80: 263–312.
- Kamp, H. and Reyle, U. 1993. *From Discourse to Logic*. Dordrecht: Kluwer.
- Lasswitz, K. 1879. Review of the *Conceptual Notation*. *Jenaer Literaturzeitung* 6: 248–9. Reprinted and translated in T. W. Bynum *Conceptual Notation and Related Articles*. Oxford: Oxford University Press, 1972.
- Quine, W. V. O. 1960. *Word and Object*. New York: Technology Press of MIT and John Wiley and Sons Inc.
- Read, S. 2012. The medieval theory of consequence. *Synthese* 187: 899–912.
- Russell, B. 1904. Letter to Victoria Welby, 2/3/1904. Reprinted in *Bertrand Russell and the Origins of Analytic Philosophy*. Edited by R. Monk and A. Palmer. Bristol: Thoemmes Press: 58–60.
- Russell, B. 1905. On denoting. *Mind* 14: 479–93.
- Russell, B. 1910–11. Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society* 11: 108–28.
- Russell, B. 1913. *Theory of Knowledge*. London: George, Allen and Unwin, 1984.
- Russell, B. 1914. Logic as the essence of philosophy. In his *Our Knowledge of the External World*. London: George Allen and Unwin: 42–69.
- Russell, B. 1919a. The philosophy of logical atomism, III. *The Monist* 29:

332–63.

Russell, B. 1919b. *Introduction to Mathematical Philosophy*. London: George Allen and Unwin.

Russell, B. 1951. Reply to criticisms. In *The Philosophy of Bertrand Russell*. Edited by P. A. Schilpp, 3<sup>rd</sup> edition. New York: Tudor Publishing Company: 681–741.

Russell, B. and Whitehead, A. N. 1910–13. *Principia Mathematica*. Cambridge: Cambridge University Press.