ON EXPLAINING THAT*

Sentences (1) and (2) below are truth-conditionally equivalent, while (2) and (3) are not:

(1) Nora explained why Fido barked

(2) Nora explained the fact that Fido barked

(3) Nora explained that Fido barked.

Sentence (3) can be used to report a situation in which Nora uttered 'Fido barked' in response to a question. (Nick asked why the burglar ran off. Nora explained that Fido barked.) In cases of this sort, *that Fido barked* is not the explanandum—the thing explained. Rather, it is what Nora said in the course of explaining something else (*e.g.*, the fact that the burglar ran off). But if (1-2) are true, *the fact that Fido barked* is what Nora explained. In this respect, 'the fact that Q' is relevantly like 'why Q' and not 'that Q'. This is somewhat puzzling, especially given the following Fregean view:

(α) the referent of 'that Q' is the proposition expressed by 'Q'; and facts are

true propositions; so if 'Q' is true, then the referent of 'that Q' is the fact that Q.

I endorse (α).[1] But even if one rejects (α), one might find it surprising that (2) and (3) have different truth conditions. For how can Nora bear the relation expressed by 'explained' to the fact that Fido barked, without bearing the same relation to the referent of 'that Fido barked'? One can't, for example, admire the poet Frost without admiring the referent of 'Frost'.

I suggest a semantics for (2-3) that preserves (α) and abandons

(β) instances of 'x **V** that Q' are true, iff

x bears the relation expressed by (the verb) '**V**' to the referent of 'that Q'.

If (3) is true, Nora bears *some* relation to the referent of 'that Fido barked', but not a relation expressed by 'explained' (at least not in the ordinary sense of 'expressed'). Drawing on a development of Davidson's event analysis, I argue that verbs like 'explain' take direct objects of two different sorts, which are associated with two distinct thematic roles. If this is correct, it is at best misleading to speak of the relation expressed by such verbs. And this has implications for 'said', 'believes', *etc*.

1. Proposal

If (1-3) are true, something *happened*. Nora *acted*. The standard reasons for adopting an event analysis also apply. For example, if x quickly explained why Q, x explained why Q. So one might say the logical forms of (1-3) are: $\exists e[\text{Explain}(e, \text{Nora}, \text{why Q})]$; $\exists e[\text{Explain}(e, \text{Nora}, \text{the fact that Fido barked})]$; and $\exists e[\text{Explain}(e, \text{Nora}, \text{that Fido barked})]$. But this makes it hard to see how (2) and (3) can differ in truth value. In particular, if π is the fact that Fido barked, and π is the referent of 'that Fido barked', then (2) and (3) are both true iff $\exists e[\text{Explain}(e, \text{Nora}, \pi)]$. Suppose, however, that the logical form of (2) is

(2L) $\exists e[\text{Agent}(e, \text{Nora}) \,\&\, \text{Explaining}(e) \,\&\, \text{Theme}(e, \text{the fact that Fido barked})]$

where the Theme of a Φ-ing is, intuitively and roughly, the thing Φ'd. By way of comparison, if Nora kicked the ball, $\exists e[\text{Agent}(e, \text{Nora}) \,\&\, \text{Kicking}(e) \,\&\, \text{Theme}(e, \text{the ball})]$. I assume that certain events have abstract Themes. And if x explained the fact that Q, this explanandum is the Theme of x's explaining.[2]

According to this kind of event analysis, a verb and its arguments are represented as separate conjuncts of a complex event predicate, and the arguments are associated with *thematic roles*. Davidson originally suggested that an action verb which otherwise appears to have *n* arguments is true of ordered $n+1$-tuples, the extra position being for a variable (ranging over events) bound by an existential quantifier; but following Castañeda, one might suspect that notation like '$\exists e[\text{Kick}(e, x, y)]$' masks relevant structure.[3] If x kicked y, it *follows* that there was a kicking, and that x did something. More importantly, a range of semantic phenomena—from entailment patterns exhibited by causative verbs to quantificational constructions involving plural noun phases—can be explained by treating action verbs as 1-place event sortals, and associating arguments with thematic relations (which hold between events and the semantic values of linguistic expressions that appear in the relevant argument positions).[4]

It may be useful to highlight the contrast between these two eventish proposals in categorical terms. On a traditional view, an intransitive verb like 'yelled' expresses a function of type $\langle i, t \rangle$, from individuals to truth values; a transitive verb like 'kicked' expresses a function of type $\langle i, \langle i, t \rangle \rangle$, from

individuals to functions from individuals to truth values. A modification of this view is that 'yelled' expresses a function of type $\langle i, \langle e, t \rangle \rangle$, from individuals to functions from events to truth values; while 'kicked' expresses a function of type $\langle i, \langle i, \langle e, t \rangle \rangle \rangle$, from individuals to functions of the sort expressed by 'yelled'. But on a thematically elaborated event analysis, one would say that 'yelled' and 'kicked' both express functions of type $\langle e, t \rangle$, from events to truth-values; although at this point, one might question the value of saying that verbs express functions, as opposed to saying that verbs are satisfied by events. For on the thematically elaborated view, the syntactic arguments of the verb do not denote possible arguments of a function expressed by the verb; 'kicked' is not even true of ordered pairs, much less ordered pairs consisting of a kicker and kickee. From this perspective, intuitions of relationality reflect the fact that 'kicked' is associated with two thematic roles—and thus indirectly associated with ordered pairs $\langle x, y \rangle$ such that x kicked y (since y is the Theme of a kicking whose Agent is x).

This suggests an account of why (2-3) differ semantically. Suppose the logical form of (3) is

(3L) $\exists$e[Agent(e, Nora) & Explaining(e) & Content(e, that Fido barked)]

where the Content of a Φ-ing is, intuitively, the meaning of the Φ-ing. I assume that some explainings have Contents, since propositions are expressed in the course of performing the relevant actions. But the Content of an explaining is not the explanandum; it is, typically, the explanans. This distinction seems directly relevant to the truth-conditional difference between (2) and (3).

We can also extend this idea to

(4) The fact that Fido saw the burglar explains the fact that Fido barked.

Suppose that 'explain' has—in addition to its eventive form, which appears in (2-3)—a *stative* form that appears in (4). Using subscripts, let 'explain$_\sigma$' be a sortal for events, while 'explain$_\sigma$' (like 'knows') is a sortal for states. One can quantify over states as in

(4L) $\exists$s[Subject(s, the fact that Fido saw the burglar) & Explaining$_\sigma$(s) &

Theme(s, the fact that Fido barked)].

3

On this view, (4) is true iff: there is a state of explaining whose subject is the fact that Fido saw the burglar, and whose Theme is the fact that Fido barked. I assume that states have Subjects (not Agents), and that to be the Subject of a state is to have a certain property. (One might also say that states have Objects instead of Themes.) But in any case, if (4L) is true, the fact that Fido barked is the explanandum. While the fact that Fido saw the burglar does not act—it does not explain$_\varepsilon$ anything—it can explain$_\sigma$ (*i.e.*, have the property of explaining$_\sigma$) the fact that Fido barked. The verb 'provide' provides a useful comparison here. If Nora built a well, she may have provided$_\varepsilon$ a source of water for someone; but a river can also provide$_\sigma$ a source of water for someone, where 'provide$_\sigma$' is a near synonym to 'be'. Similarly, Nora can provide$_\varepsilon$ an answer to a question; while the fact that Fido barked can provide$_\sigma$ an answer.

2. Some Implications

Setting aside for now these last remarks about states, I have claimed that the logical forms of (2-3) are:

(2L) $\exists e[\text{Agent}(e, \text{Nora}) \ \& \ \text{Explaining}_\varepsilon(e) \ \& \ \text{Theme}(e, \text{the fact that Fido barked})]$

(3L) $\exists e[\text{Agent}(e, \text{Nora}) \ \& \ \text{Explaining}_\varepsilon(e) \ \& \ \text{Content}(e, \text{that Fido barked})]$

where the Theme of an explaining—the explanandum—is distinct from the Content of an explaining. If this is correct, one can grant that (2) and (3) have different truth conditions, while retaining

(α) the referent of 'that Q' is the proposition expressed by 'Q'; and facts are

true propositions; so if 'Q' is true, then the referent of 'that Q' is the fact that Q.

One can say that the referent of 'that Fido barked' is the proposition expressed by 'Fido barked', and this (true) proposition *is* the fact that Fido barked. On this view, the difference between (2) and (3) lies not in the direct objects having different referents, but in the different thematic relations to a single proposition. A proposition can be the Theme *but not the Content*, or the Content *but not the Theme,* of an explaining.

On the other hand, if the proposed treatment of 'explain' is correct, we must reassess

(β) instances of the form 'x **V** that Q' are true, iff

x bears the relation expressed by (the verb) '**V**' to the referent of 'that Q'.

4

For (3) is an instance of 'x **V** that Q'. But it is at least misleading to speak of *the relation expressed* by 'explained$_\varepsilon$'. In saying that a word *expresses* a relation, one implies that the word *is true of* (ordered) n-tuples of things. On the thematically elaborated event analysis, however, verbs are true of events or states (not ordered n-tuples). More importantly, there is no *unique* thematic relation associated with 'explained$_\varepsilon$'; we can speak of an Agent-Theme relation or an Agent-Content relation. And if either relation is primary, it is intuitively the Agent-Theme relation, which holds between an explainer and the thing explained (the explanandum). But in 'x V that Q', the phrase 'that Q' specifies a Content.

One can preserve (β) by adopting an ambiguity hypothesis: 'explains$_{\varepsilon1}$' is associated with an Agent-Theme relation, while 'explains$_{\varepsilon2}$' is associated with an Agent-Content relation. But other things equal, one wants to avoid such hypotheses and the meaning postulates that inevitably accompany them; and at least prima facie, the same matrix verb appears in (2) and (3). Alternatively, one might deny that (3) has the form described by (β). If x explained that Q, x did so *by* doing something else—*e.g.*, asserting that Q in response to a question. So perhaps the surface form of (3) masks a more complicated logical form like the following:

$$\exists e \exists x \exists f[\text{Agent}(e, \text{Nora}) \ \& \ \text{Explaining}_\varepsilon(e) \ \& \ \text{Theme}(e, x) \ \& \ \text{By}(e,f) \ \&$$

$$\text{Agent}(f, \text{Nora}) \ \& \ \text{Asserting}(f) \ \& \ \text{Content}(f, \text{that Fido barked})].$$

On this view, 'that Fido barked' is an argument of an unspoken verb, which belongs to an adverbial phrase. But one needs independent evidence for the unspoken material, and an account of why

(15) Nora explained that Fido barked on Tuesday

is not *three* ways ambiguous. Prima facie, the adverbial phrase 'on Tuesday' should be able to modify all three verbs: 'explain', 'bark', and the unvoiced 'assert'. But (15) cannot mean that Nora explained something by *asserting on Tuesday* that Fido barked. Nor can

(16) On Monday, Nora explained on Tuesday that Fido barked

mean that: Nora explained something on Monday, by asserting on Tuesday that Fido barked.[5]

5

If we reject (β), this has implications for sentences like

(17) Nora said that Fido barked

(18) Nora believes that Fido barked

whose logical forms would be:

(17L) ∃e[Agent(e, Nora) & Saying(e) & Content(e, that Fido barked)]

(18L) ∃s[Subject(s, Nora) & Believing(s) & Content(s, that Fido barked)].

If (3) and (17) have the same logical form, we should pause before claiming that (17) is true iff Nora

bears the saying-relation to the referent of 'that Fido barked'. Similarly, if (18) is true, Nora is the subject

of a state *whose Content is* that Fido barked. But if 'believe' is not true of Subject-Theme pairs, we need

not think of believing as something *done to* a proposition.[6] (Compare 'grasp the proposition', which

suggests a Theme of the grasping.) So we should be wary of treating (17-18) on the model of

(19) She kicked it.

For (17-19) may not be instances of the common logical form 'Rxy', unless 'R' stands for the

(disjunctive) relation: x is the Agent-or-Subject of an event-or-state whose Theme-or-Content is y.

Fregeans should welcome this consequence: believing that P need not be construed as "mentally

grabbing" and affirming a thought. But paratactic theorists face a serious difficulty. According to

Davidson, the truth conditions of (17) are given by the second sentence in

(17P) Fido barked. Nora said that.

Here, 'that' is used to demonstrate (an utterance of) the first sentence.[7] One might extend this idea to

(1) Nora explained why Fido barked.

(1P) Fido barked. Nora explained why.

For the second sentence in (1P) arguably captures the truth conditions of (1), so long as the reference of

'why' in (1P) is fixed by the first sentence in (1P). But this analogy to (17) breaks down with respect to

(3) Nora explained that Fido barked

(3P) Fido barked. Nora explained that.

For (3P) captures the truth-conditions of (1), which is true iff Nora's *explanandum* was (the fact) that Fido barked; (3P) does not capture the truth-conditions of (3), which is true iff Nora's *explanans* was (the fact) that Fido barked.

Correlatively, the demonstrative in (3P) is associated with a Theme. So the second sentence in (3P) is true iff ∃e[Agent(e, Nora) & Explaining(e) & Theme(e, Δ)]; where Δ is the item demonstrated with 'that'. For Davidson, Δ will be an utterance of 'Fido barked'; although a paratactic theorist willing to posit facts—if only to avoid the consequence that explananda are *utterances*—could allow that Δ is the fact that Fido barked. But on any view, if the second sentence in (3P) is true, the referent of 'that' is the thing explained (and not the Content of Nora's speech act). Hence, the second sentence in (3P) has *different* truth conditions than (3), in which 'that Fido barked' is associated with a Content. In short, paratactic theorists need to express the truth conditions of (3) as follows: ∃e[Agent(e, Nora) & Explaining(e) & Content(e, Δ)]. But this is just what (3P) fails to deliver.

3. Conclusion

The logical forms of (2) and (3) are, respectively, (2L) and (3L):

(2) Nora explained the fact that Fido barked

(2L) ∃e[Agent(e, Nora) & Explaining(e) & Theme(e, the fact that Fido barked)]

(3) Nora explained that Fido barked

(3L) ∃e[Agent(e, Nora) & Explaining(e) & Content(e, that Fido barked)].

The semantics of 'that'-clauses should be pursued within a thematically elaborated event analysis that makes a distinction between Themes and Contents. And this suggests that it may be time to retire

(β) instances of 'x **V** that Q' are true, iff

x bears the relation expressed by (the verb) '**V**' to the referent of 'that Q'.

*For helpful comments, thanks to: Mark Baker, Susan Dwyer, Ernie Lepore, and James McGilvray.

1. For defense, see Paul Pietroski, *Causing Actions* (Oxford: OUP, 2000). Here and throughout, I suppress relativization to context for simplicity.

2. If x is the Agent of an event whose Theme is an abstract object y, then perhaps x manipulates some concrete *representation of* y. But I take no stand on this issue here. I allow for events that have multiple Themes (and Agents), as in 'Three scientists explained six facts'; see Barry Taylor, *Modes of Occurrence* (Oxford: Blackwell, 1985) and Barry Schein, *Plurals* (Cambridge, MA: MIT, 1993). But for simplicity, I assume that the events of explaining in question have unique Themes. This is *not*, however, to assume that *facts* have unique explanations. If a single fact can be twice explained, so be it.

3. Donald Davidson, "The Logical Form of Action Sentences" and Hector-Neri Castañeda, "Comments," both in Nicholas Rescher, ed., *The Logic of Decision and Action* (Pittsburgh: University of Pittsburgh Press, 1967). Davidson later endorsed thematic separation; see his "Adverbs of Action," in B. Vermazen and M. Hintikka, eds., *Essays on Davidson: Actions and Events* (Oxford: Clarendon Press, 1985).

4. For example, see Terence Parsons, *Events in the Semantics of English* (Cambridge, MA: MIT, 1990); Schein, *op. cit.*, who argues persuasively that certain plural constructions require a thematically elaborated event analysis; Pietroski, *op. cit.*; Elena Herburger, *In the Event of Focus* (Cambridge, MA: MIT, forthcoming); Paul Pietroski, *Events and Semantic Architecture* (Oxford: OUP, forthcoming).

5. Cf. Jerry Fodor, "Three Reasons for not Deriving 'Kill' from 'Cause to Die'," *Linguistic Inquiry*, 1 (1970): 429-38.

6. In this regard, consider "measurement analogies" for belief ascriptions—as discussed in Robert Matthews, "The Measure of Mind," *Mind*, 103 (1994): 131-46.

7. Donald Davidson, "On Saying That," *Synthese*, 19 (1968): 130-46. For more recent discussion, see Herman Cappelen and Ernest Lepore, "Varieties of Quotation," *Mind*, 106 (1997): 429-50.