

David Beaver et al. (eds.), *The Construction of Meaning* (CSLI, 2002)

I Wonder What Kind of Construction That This Example Illustrates

ARNOLD M. ZWICKY

1 The phenomenon

In modern English, WH constructions (whether of the interrogative or the relative type, whether in a main or a subordinate clause) have a particularly simple form: an initial XP containing a WH word (XP[...WH]), followed by a clause missing an XP (S/XP):

(A) Subordinate interrogative:

I want to tell you
[[what experiences] [I've had ___ here in my work.]]
XP[... WH] S/XP

(B) Main interrogative:

[What experiences] [I have I had ___ here in my work?]
XP [... WH] S/XP

(C) Subordinate relative:

I want to tell you about the experiences
[[which] [I've had ___ here in my work.]]
XP [... WH] S/XP

However, on occasion you can hear, or read, subordinate interrogatives with an additional ingredient, the subordinator *that* following

*Thanks to those who have offered comments, data, or references: among them, Ash Asudeh, David Beaver, Emily Bender, Paul Kiparsky, Geoffrey Pullum, Elizabeth Traugott, Louise Vigeant.

The Construction of Meaning
David Beaver, Luis D. Casillas Martínez, Brady Z. Clark and Stefan Kaufmann. (eds.)
Copyright ©2002, CSLI Publications

the initial XP; for instance, corresponding to (A) above there is the attested sentence:

I want to tell you
 [[what experiences] [that] [I've had ___ here in my work.]]

I'll refer to these as 'WH+*that* clauses'.

Most informants judge WH+*that* clauses to be simply ungrammatical (though comprehensible, and perhaps archaic, perhaps colloquial, in tone). Nevertheless, examples do occur with modest frequency, in the speech of both British and American speakers who are not notably non-standard (or inclined to deliberate archaism). For at least some of these speakers, it seems clear that WH+*that* clauses are not speech errors; one of my sources, who supplied three of the 27 attestations below, seems to use these clauses fairly frequently in his radio broadcasts, and another, who produced example (18) below, gave me positive grammaticality judgments on a number of other examples.

The attested examples below were collected fortuitously, in the course of ordinary listening and reading, a fact that might account for the rather small number of examples on the list; since WH+*that* clauses are comprehensible, listeners and readers might easily fail to notice the (to them) intrusive *that*. It's also true that, at least as far as I'm concerned, experience with WH+*that* clauses tends to make them unremarkable, hard to detect. Finally, for reasons that will soon become clear, WH+*that* clauses are probably not enormously frequent, even for speakers for whom they are ordinary.

Items (1)–(5) on the list are from Radford (1988:500), item (6) from Radford (1988:501). The speakers are all, I believe, British or Irish, and the identifications are Radford's. The remaining 21 examples are my own. The speakers are all American, though from various regions and ethnic/racial groups, and adults, though from several generations. I have identified by name speakers who are public personalities and those who were quoted in the media, but have removed such identifications for all other sources. (Of course, there were a number of speakers whose names I never caught in the first place.)

In what follows I will refer to these examples as the 'RZ corpus' (collected by Radford/Zwicky); later I will discuss another collection of examples that has quite recently come to my attention, the 'ST corpus' of Seppänen & Trotta (2000), which has 90 examples from sources comprising roughly 150 million words.

1.1 Examples

- (1) I'm not sure what kind of a ban that FIFA has in mind
(Bert Millichip, BBC radio 4)
- (2) We'll see what sort of pace that Daley Thompson's running at
(Ron Pickering, BBC 1 TV)
- (3) It'll probably be evident from the field which of the players that
are feeling the heat most
(Jimmy Hill, BBC 1 TV)
- (4) ... no matter what choice that the committee makes
(Bob Geldof, BBC 1 TV)
- (5) We can look at our statistics and see what sort of pattern that
we get
(Bob Morris Jones 'at a Linguistics conference recently')
- (6) What a mine of useless information that I am!
(Terry Wogan, recorded on BBC Radio 2)
- (7) Regardless of which version of the FEC bill that is passed, ...
(authority being interviewed on NPR's 'All Things Considered',
8/31/94)
- (8) However many people that were there, ...
(Maxx Faulkner on WCBE, 1994)
- (9) I hadn't realized just how many people that were there.
(Maxx Faulkner on WCBE, 1994)
- (10) I want to tell you what experiences that I've had here in my work.
(Columbus Stonewall presentation, 9/1/94)
- (11) ... people are basing [i.e. judging] professors on what kinds of
grades that they expect to receive from them.
(audience comment on NPR's 'Talk of the Nation', 9/8/94)
- (12) ... no matter what kind of people that they are dealing with ...
(disc. of Waco incident, NPR's 'Morning Edition', 7/20/95)
- (13) Unless you know how much water that you need to drink, ...
('Morning Edition' interview, 7/21/95)

- (14) [verb of speaking, not caught] . . . about careers, and about what kinds of things that they [the children] could do . . .
(interviewee re Africentric schools in Columbus, WCBE, 3/29/96)
- (15) It all depends on how much work that you had to do.
(astronaut Shannon Lucid, interviewed on NPR, 9/16/96)
- (16) What this is about is our providing them with whatever assistance that they need.
(Monitor Radio [heard on WCBE] quote from Florida state official, 12/4/96)
- (17) Look at it as a tribute to how deeply that I feel about you.
(TV show 'Bob', Bob Newhart character speaking)
- (18) We don't know to what degree that the dialects are converging.
(1997 speaker at Ohio State University)
- (19) . . . until late in the week, when we see how many people that were arrested . . .
(interviewee on NPR's 'All Things Considered', 5/26/98)
- (20) . . . maybe you could share with us what information that you have . . .
(Maxx Faulkner on WCBE, 6/11/98)
- (21) "I am pleased and frankly surprised at how soon after the hearing that the judge approved it," said Mary Stowell . . .
(New York Times story, p. B2, 7/25/98)
- (22) Whatever actions that we do need to do, we will . . .
(interviewee on WCBE re tobacco restrictions, 8/20/98)
- (23) . . . we asked what sort of health care that they rely on.
(interviewee on KQED-FM, 9/2/98)
- (24) I'd like to make whatever contribution that I can.
(public radio interviewee, 9/14/99)
- (25) Consumers don't realize how many scrips that they [pharmacists] have to fill in a day.
(interviewee on NPR's 'Talk of the Nation', 12/15/99)
- (26) I was surprised at . . . how few things that they needed [in India] to have to have fun.
(interviewee on public radio program, heard on KALW, 4/19/00)

- (27) Yes, I fully understand what uproar that would come about.
(public radio interview of Miami Chief of Police, 5/2/00)

To these examples from spoken English, we can add two examples from standard written English, collected by Seppänen (1994):

- (28) A number of semantic, formal and functional criteria have been provided in the literature, and definitions vary as to which of these types of criteria that are used.
(member of the University of Göteborg English Department)
- (29) What little hostility that remained against him and his men among the villagers disappeared.
(John Fowles, *The Magus*)

I will include these in the RZ corpus.

2 Observations

The examples in the inventory have several common characteristics, which I now describe.

2.1 Interrogative, not relative

English WH constructions fall into two large classes, interrogative and relative, with many subtypes of each, plus some that show characteristics of both. A collection of these characteristics is assembled in Zwicky (1986).

The WH+*that* clauses in the inventory above are all of the interrogative type, though of several different subtypes. Certainly, none of them is of a transparently relative type;¹ that is, there are no examples of ordinary restrictive relatives with WH+*that*² (like *–the people whose house*

¹Radford (1988:486) gives one example of an apparent WH+*that* in a (restrictive) relative clause: *England put themselves in a position whereby that they took a lot of credit for tonight's game* (Ron Greenwood, BBC Radio 4). The appearance of the formal *whereby* in this otherwise colloquial bit of sports reporting is suspicious, as is the fact that there are no relative clause examples in the RZ corpus, and no example there with a one-word WH expression (Section 2.4 below), not to mention the fact that informants unhesitatingly reject such examples. My guess is that this is a production error, an on-line blend of a formal, 'fancy' construction with *whereby* and a colloquial, everyday construction with *that* (... *a position that they took a lot of credit for tonight's game from*, say). For further discussion, see Section 5.1.

²From here on, a leading hyphen marks an example of a type which does not occur in the data and which I predict would be judged as ungrammatical even by speakers who accept WH+*that*.

that I visited) or of ordinary appositive relatives (like *–Kim and Terry, whose house that we visited; –I suggested we should leave, which idea that everyone approved*). There are also no examples of free relatives with WH+*that* (for example, *–What that I had in my hand exploded* or *–What thing that I had in my hand exploded*, parallel to *What I had in my hand exploded*), but this gap is explicable on other grounds; see 2.4 below.

In any case, the attested sentences exemplify several different interrogative-type constructions, including indirect questions, main and subordinate exclamations, concessive arguments, and concessive modifiers:

• **Indirect questions:**

- Objects of verbs of communication: *tell someone* (10); [*speak about* (14); *share with someone* (20); *ask* (23).
- Objects of verbs of mental action: *be sure* (1); *see* ‘understand’ (2), (5), (19); *realize* (9), (25); *know* (13), (18); *understand* (27).
- Other objects: *base* ‘judge’ *someone on* (11), *depend on* (15), *vary as to* (28).
- (Postposed) subject of verb of mental action: *be evident* (3).

• **Exclamations:**

- Main exclamation (6).
- Objects of exclamatory verbs: *be surprised at* (6), (26); *be a tribute to* (17).

• **Concessives:**

- WH-*ever* arguments (16), (24).
- WH + concessive *little* (29).
- Sentence modifiers: WH-*ever* (8), (22); *no matter* (4), (12); *regardless of* (7).

2.2 Finite only

All the attested sentences have WH+*that* in finite clauses. There are several types of non-finite interrogative WH constructions: the main-clause infinitival question (*Which book to read? Whose paper to look at?*), the indirect infinitival question (*I wonder which book to read; I don't know whose paper to look at*), and the main-clause infinitival exclamation (*What a question (for Terry) to ask!*), but WH+*that* seems utterly impossible in these contexts: *–Which book that to read? –Whose paper that to look at? –I wonder which book that to read. –I don't know whose paper that to look at. –What a question that (for Terry) to ask!*

2.3 Not inverted (hence mostly subordinate)

The one main-clause example, (6), is an exclamation, and therefore not inverted. Main-clause questions, which are inverted, absolutely reject that: *-What sort of a pace that is Daley Thompson running at?* (cf. (2)), *-How much water that do you need to drink?* (cf. (13)).³ Even main-clause subject questions, which superficially do not appear to be inverted, absolutely reject *that*, and this is so whether the clause has an auxiliary to invert or not: *-What sort of person that would talk like that?* *-What sort of person that talks like that?*

(Many would analyze such examples as involving inversion, the effects of which are masked by the fronting of the subject XP[... WH]. There is an enormous literature on the matter, going back to the earliest days of generative syntax, with various versions of the double-movement analysis, not all of them actually involving 'movement', confronting various versions of the no-movement analysis. The facts about WH+*that* clauses would at first glance appear to support some sort of double-movement analysis. But see Section 3.2.4 below.)

There are a few types of inverted subordinate constructions (*Kim was more pleased than was Sandy*), but these either don't involve WH or else mirror main-clause inversion directly (*I wondered which candidates did they prefer*).⁴ In any case, it would seem that inversion is sufficient to block WH+*that*, and that subordination is not necessary.

But let's not be so hasty here. The entire discussion rests on a single example, (6) *What a mine of useless information that I am!* I have been treating this example as unequivocally acceptable for speakers who also accept structures that are attested in multiple examples from the RZ corpus. This might be correct (for at least some speakers), though I would be happier if I actually had some informants whose judgments indicated a grammar in which both embedded WH interrogative-type constructions and main-clause WH exclamations had alternatives with *that* (while main-clause WH questions lacked such alternatives). But I am inclined to believe that example (6) should not be lumped together with the other 26 for the purposes of analysis. There are three alternatives to treating (6) as a straightforward instance of a WH+*that* clause, and I favor the third of these.

³Henry (1995:108) notes that in Belfast English, where (pg. 107) examples like *I wonder which dish that they picked* and *They didn't know which model that we had discussed* are grammatical, main-clause questions absolutely reject *that*: **Which dish that they picked?*/**Which dish that did they pick?*

⁴Henry (1995:108) notes that in Belfast English such examples are ungrammatical: **I wondered which dish that did they pick?*

(a) The first alternative interpretation of (6) is that it is a production error. One possible story would be that the speaker embarked on a rather long fronted NP (beginning with the WH modifier *what*), holding in store a clause missing an NP of this type (a S/NP), and that by the time he reached the end of the NP, where the expression with the significant content (*useless information*) is located, he lost track of some of the earlier parsing and treated this expression as a head NP in combination with a following modifying relative clause, a different type of S/NP—and then supplied an explicit marker, *that*, for this relative clause. On this interpretation, (6) is a one-time, on-line blend of a main-clause WH exclamation and a *that* relative.

This is not at all implausible. Speakers do blend constructions in this way, beginning on one track and then shifting partway through onto another, overlapping, track. In fact, this is essentially the story I'll be telling about the historical origin of systematic WH+*that* clauses for some speakers of English. Of course, when we hear a single example, we can't tell whether it's a production error or part of the speaker's linguistic system. So it is that one person's hypercorrection becomes an instance of another person's grammatical rule; *They told it to Kim and I* could be either, though these days it's almost surely the latter.

In any case, it's possible that (6) doesn't need to be figured into an analysis of WH+*that* clauses in English at all.

(b) The second alternative interpretation of (6) is that it's definitely in its speaker's variety of English—in fact, it's in my variety of English—but as an exclamatory NP, not an exclamatory clause. Exclamatory NPs like *What a good dog!*, *What clever students!*, and *What an idiot!* are in fact reasonably common, and indisputably grammatical. They can also have postnominal modifiers: *What a good dog in that house!* *What clever students taking that exam!* These postnominal modifiers can be relative clauses (either *that* relatives or WH relatives, though relative clauses with *which* tend to be dispreferred in favor of those with *that*): *What a good dog that's in that house!* ? *What a good dog which is in that house!* *What clever students that/who are taking that exam!*⁵

⁵ A similar use of WH question expressions leads to (only apparent) WH+*that* clauses in standard English. *Which clever students?* (a NP) is grammatical as a free-standing expression in a context where clever students have been referred to. Such WH questions NPs can have postnominal modifiers (*Which clever students in your class?*), including postnominal relatives (*Which clever students that/who are in your class?*), even zero relatives (*Which clever students you graded hard?*). NPs like *Which clever students that are in your class?* then look superficially like WH+*that* clauses, but are not clauses at all.

Given all this, exclamatory WH expressions like *What a good dog you are!* could be either clauses or else NPs with postnominal zero relatives. There are exclamatory expressions that are unambiguously NPs—*What a good dog (in that house)!*—and exclamatory expressions that are unambiguously clauses—*What wonderful people met me at the airport! How beautiful you are!*—but those with fronted non-subject NPs could be either. And example (6), *What a mine of useless information that I am!*, is one of these. Not that I know how to discover which of these two, semantically very close, structures a speaker intended in uttering a particular expression.

(c) Yet another interpretation of (6) depends on knowing a significant fact about the person who uttered it. Terry Wogan (I am told by David Beaver) is not only a speaker of Irish English, but a proud speaker of this variety, given to exaggerating his Irishness. We might then dismiss this example as nothing more than pseudo-Irish archaism.

However, I must point out that the ST corpus includes 6 (of its 90) WH+*that* examples that Seppänen & Trotta (2000:171) classify as ‘exclamative’, like (6) above. Unfortunately, their article doesn’t cite any of these examples, nor are the speakers/writers of them identified as to regional dialect. But these items in the ST corpus make me disinclined to dismiss (6), and inclined to view it, tentatively, as grammatical for (some varieties of) Irish English, where it would serve as the WH parallel to *it*-exclamatives like *It’s a mine of information that I am!* That is, (6) would be an instance of a WH construction distinct from the interrogative construction types illustrated by the other RZ examples.

In any case, I will not be discussing example (6) further here.

2.4 The two-word minimum and the Lexical Head Restriction

The most striking shared characteristic of the attested examples is that in each, the XP[... WH] has more than one word in it.

There are no attested examples like *–I’m not sure what that FIFA has in mind* (cf. (1)), *–As for the players, it’ll probably be evident from the field which that are feeling the heat most* (cf. (3)), *–Whatever that we do need to do, we will* (cf. (22)), *–I’m not sure who that finished last* (cf. the significantly better *I’m not sure which contestant that finished last*), or *–I know whose that finished last* (cf. the significantly better *I know whose horse that finished last*). This is not just a pattern in 29 examples; the person who was the source of example (18) judged all

examples that had a single-word XP to be absolutely ungrammatical, while finding most of a sampling of sentences from the RZ corpus to be grammatical. Other speakers find none of the RZ examples entirely acceptable but judge their single-word XP counterparts to be much worse, and I have found no one who makes the opposite judgments, that is, who prefers single-word XPs to longer ones.

The unacceptability of single-word XP[... WH]'s in WH+*that* clauses means that several WH constructions cannot have variants with *that*—because these constructions require the fronted XPs in them to be single words. This is the case for both free relatives and for the subjects of pseudocleft sentences (but not for indirect questions).

- **Indirect question:**

I don't know [where these vases came from].

I don't know [which city these vases came from].

I don't know [from where these vases came].

- **Pseudocleft:**

[Where these vases came from] is Persepolis.

[*Which city these vases came from] is Persepolis.

[*From which city these vases came] is Persepolis.

- **Free relative:**

I have visited [where these vases came from].

I have visited [*which city these vases came from].

I have visited [*from where these vases came].

Now, the pseudocleft construction is certainly of the interrogative, rather than relative, type, but it has no variant with *that*, because of an irresolvable conflict of restrictions in sentences like –*Where that these vases came from is Persepolis*; the pseudocleft construction requires a single-word fronted XP, while the WH+*that* construction requires a multi-word fronted XP. Similarly with the free relative construction. Even if WH+*that* were possible with some relative-type constructions, free relatives could not have a variant with *that*, because of the irresolvable conflict of restrictions in sentences like –*I have visited where that these vases came from*.

The condition on WH+*that* clauses is not, however, really a matter of multi-word vs. single-word XPs. –*I know from what that you took it* (with the two-word XP *from what*) is no improvement over –*I know what that you took it from* (with the one-word XP *what*), but both contrast with the acceptable, or at least more acceptable, *I know from what box that you took it*. It takes a lexical (not grammatical) word, like

the noun *box*, and in addition a preposition won't do (as we saw with *-I know from what that you took it*), though a noun or adverb will.

The minimal XP[... WH] then has two words, one of them a WH word, the other a lexical word, of category N (the *contribution* of *whatever contribution* in (24)) or Adv (the *deeply* of *how deeply* in (17)); the WH word is a modifier of the lexical word.

In more complex XPs, the pattern of WH modifier plus lexical N or Adv head is repeated, but with elaborations.

In most of the examples the modified head word is a noun (*people, things, pace, experiences, water, information, health care*, etc.); in two examples, it is an adverb (*deeply* in (17), *soon* in *how soon after the hearing* in (21)). My informant who was the source of example (18) and who found the noun examples generally acceptable, nevertheless rejected the adverb example (17)—this informant was not given (21)—and found two constructed adverb examples (*no matter how long that we were in class...* and *I don't know how long that they were in China*) to be at best borderline. I take this as evidence that some speakers have a more constrained grammar, in which the modified head word must be from the category N, period. Put another way: if you accept WH+*that* clauses at all, you accept those with a lexical N as the modified head word.

The attested examples don't include any where the modified head word is an adjective, but I would venture that those who accept adverb examples like (17), with *how deeply*, and (21), with *how soon*, will also accept parallel adjective examples like *Look at it as a tribute to how deep that my feelings about you are*.

I am less sure about the acceptability of invented examples with headless XPs that nevertheless have a noun-like modifier as remainder: a quantifier (as in *I realized how much/many that I had seen*) or a sortal noun (as in *I wasn't sure what kind/sort that I had bought*). The question is whether such words are sufficiently 'lexical'. I am similarly unsure about invented examples with the indefinite pro-N *one*, like *They couldn't be sure which one that they should choose*.

Turning now to the WH modifiers in the RZ corpus, they pretty well cover the full range of types in modern English:

- *what kind/sort of* N_{SG} (1), (2), (5), (12), (23);
what kinds of N_{PL} (11), (14)
- *which of* NP_{PL} (3), (28)
- *what(ever)/which* N_{SG} (4), (7), (16), (18), (20), (24), (27);
what(ever) N_{PL} (10), (22)

- *how much* N_{SG} (13), (15); *how(ever) many* N_{PL} (8), (9), (19), (25); *how few* N_{PL} (26)
- *what little* N_{SG} (29)
- *how Adv* (17), (21)

Missing, by accident I assume, are many types of modifiers that are longer and more complex than these: *how much/many of NP* (*I don't know how much of the sugar that was in the canister, I don't know how many of the marbles that were in the bag*), *how very A* (*Look on it as a tribute to how very deeply that I feel about you*), *which/what N PP* (*I don't know which marbles from the new set that were in the bag*), etc. The one short, simple modifier that's missing from the data is *whose*, as in *I don't know whose marbles that were in the bag*; I take this to be an accidental gap as well.

To summarize: a WH+*that* clause is subject to the Lexical Head Restriction (LHR):

LHR: In the XP[... WH], the WH word is (part of) a modifier of a lexical (not grammatical) word from the category N (or A).

It should now be clear that one reason WH+*that* clauses are not very frequent (for speakers who find them grammatical) is that they are subject to a number of restrictions, some of which oblige them to be fairly complex. Interrogative WH constructions with single-word fronted XPs (*I know what you did; no matter who you saw, ...*; etc.) are very common indeed (in the corpora Seppänen & Trotta used, 84% of the 12,831 WH phrases, of all types, were single-word phrases); but these are not eligible to have variants with *that*.

2.5 Optionality

So far as I can tell, double marking of subordination, with both WH and *that*, is never obligatory; WH+*that* clauses are always alternatives to plain WH clauses. The speaker of (15), *It all depends on how much work that you had to do*, could have chosen to say instead, *It all depends on how much work you had to do*. This alternation of *that* and zero occurs also in complement clauses (*I know (that) pigs can't fly*) and relative clauses (*the work (that) we had to do*).

The optionality of *that* in combination with XP[... WH] provides yet another reason for the infrequency of WH+*that* clauses. Even in uninverted finite interrogative constructions where a WH modifier has a lexical head, double marking is not guaranteed; a plain WH clause is always available.

2.6 No restriction on syntactic function

One way in which WH+*that* clauses are not restricted (given that other restrictions can be satisfied) is that the fronted XP is essentially free with respect to its syntactic function within its clause. In about half the attested examples, the fronted XP is a direct object within its clause: *what experiences that I've had* __ *here in my work* (10), for instance. In five examples, it is the object of a preposition, either stranded (*what kind of people that they are dealing with* __ (12)) or pied-piped (*to what degree that the dialects are converging* __ (18)). In two, it is an adverbial (*how soon after the hearing that the judge approved* __ (21)).

And in eight examples—(3), (7), (8), (9), (19), (27), (28), and (29)—the fronted XP is the subject of its clause: *however many people that* __ *were there, . . .* (8). Note that in clauses like *however many people were there*, the subject *however many people* appears to be located in subject position, with no 'movement' to the left, but in the WH-subject examples like (8), the *that* intervenes between this subject and its predicate, indicating that the subject has been, in some sense, displaced.

None of the attested examples has a long-distance fronting, like *I'm not sure what kind of a ban that Kim thinks FIFA has* __ *in mind* (cf. (1)). Surely this is an accidental gap in the data, the result of the low frequency of WH+*that* clauses (for speakers who have them at all) interacting with the low frequency of long-distance fronting (for modern English speakers as a group).

3 Towards an analysis

So far as I know, WH+*that* clauses in modern English have been systematically treated in the theoretical literature only by Radford (1988), who relates them (pg. 500f.) to superficially similar interrogative-type phenomena (including direct questions, indirect questions, and exclamations) in a number of other languages (Old English, Middle English, Bavarian, Norwegian, Flemish, Popular French, Colloquial Moroccan Arabic, Frisian, Irish, and Italian) and to relative clause phenomena (pg. 486) in an overlapping set of languages (Old English, Middle English, Canadian French, Dutch), to which we can add the well-known co-occurrence of subordinating conjunctions with complementizers in earlier English and in many other languages (giving sequences that translate as *if that, when that, before that, because that*, and the like).

3.1 Lack of parallelism between the modern English phenomenon and WH+*that* in other languages

These are seductive parallels, but I maintain that they are thoroughly misleading; modern English WH+*that* clauses are quite unlike their apparent forebears from 500-1000 years ago, and quite unlike their easy-to-find look-alikes in other languages, Indo-European or otherwise.

Not structure-general

My first reason for thinking that modern English WH+*that* clauses are different from these apparent parallels is that the modern English phenomenon is restricted to interrogative-type constructions; it does not occur in relative clauses or complement clauses. That is, it is restricted to a small class of constructions, rather than being 'structure-general', as would be predicted by Radford's analysis (and by many other analyses).

Abstracting away from a very large number of details,⁶ one prominent style of analysis of complementation posits three levels of clause structure, labeled in the template below (from the inside out) Z (which is S/XP or a full S), Y, and X. The upper two levels can have initial markers of subordination, labeled (from the inside out) ② and ①.

The subordination template: [X ① [Y ② Z]]

The inner subordinator ② is a neutral complementizer, like that in English. The outer subordinator ① is a fronted WH expression (interrogative or relative) or a subordinating conjunction. Depending on the construction, either or both of these positions can be unfilled; both are unfilled in unmarked complements (*I know pigs can't fly*, alongside the marked complement in *I know that pigs can't fly*, with position ② filled) and zero relatives (*the pigs I saw*, alongside the *that* relative in *the pigs that I saw*, with position ② filled, and the WH relative in *the pigs which I saw*, with position ① filled).

In a construction-based framework, facts like these, about which positions are filled by which sorts of expressions, are simply stipulated.

⁶These details include, but are not limited to: (a) the categories associated with X, Y, Z, ①, and ②; (b) whether material fills positions ① and ② by movement (in a derivational account) or by a static requirement; (c) whether the positions ① and ② are ordered with respect to one another by virtue of their linear position within X (on the left edge, immediately preceding Z), their dominating category, their sister category, or some combination of these; and (d) whether unfilled positions simply have nothing in them or are filled by a zero element. Still another variable is whether there are in fact three levels of structure, or only two.

The sharing of conditions among different constructions can be captured as well (by a hierarchical typing of constructions, as in Sag (1997), or by having one construction invoke others, as in Zwicky (1994)), but it is not forced by other assumptions. In such a framework, most varieties of modern English lack doubly marked subordination—WH+*that* interrogatives, WH+*that* relatives, Conj+*that* adverbial clauses—because no construction calls for it. It would be no surprise to discover a variety in which one particular construction was doubly marked, or one in which all constructions of a particular type (say, all WH interrogative constructions) were doubly marked, or various other possibilities.

In a principles-based framework (such as the one developed in Radford's book), an inventory of possible conditions on the occurrence of various elements (like XP[... WH], *that*, or Subject) is provided by a universal theory, individual languages make a selection from this inventory, and the interactions between these conditions are regulated primarily, if not entirely, by general principles. In particular, one condition requires that a XP[... WH] be fronted to position ①, another that elements like *that* fill position ②. If we say no more, then in a language with both conditions, doubly marked subordination is predicted for all types of WH constructions, as well as with subordinate conjunctions.

This, of course, is not standard modern English, for which constructions, WH or not, in which position ① is filled are all incompatible with *that* in position ②. A language-particular stipulation must ensure this somehow. Which of various possibilities is best is not my concern here. What's important is that if restriction is lifted, we should get the full range of doubly marked subordinate types.

This is not what we find. Only certain classes of doubly marked constructions are possible; the combinability of WH with *that* is construction-specific, not structure-general. (It's not my point here, but I very much doubt that double marking is structure-general in all, or even most, of the languages Radford cites, or might have cited.)

The LHR

Let's suppose that we have managed to incorporate the restriction to interrogative-type constructions while still allowing XP[... WH] in position ① to co-occur with *that* in position ②. We still make the wrong predictions about modern English.

The problem is that languages with doubly marked constructions almost always allow single-word WH expressions in position ①. Not only are single-word expressions allowed, they are the favored and most

frequent fillers for position ①. Almost all of Radford's examples (from languages other than modern English) have single-word fillers, in fact, and when I ask linguist colleagues to supply WH+*that* examples in languages they know, their responses invariably involve equivalents to *who that* and *which that*. These are exactly the sorts of expressions that the LHR rules out, but they are the prototypical examples of doubly marked constructions in nearly all languages that have them.

I conclude that whatever allows WH+*that* clauses for some speakers of English (in which the XPs must have lexical heads for their WH head words), it's not at all the same thing that allows the usual sort of doubly marked WH constructions (which typically have single-word XPs). In turn, this means that the modern English clauses are not continuations of the doubly marked constructions of older forms of English, but are fresh developments.

Optionality

It might seem that the optionality of *that* in WH clauses (Section 2.5 above) would have a straightforward account in Radford's approach: in the standard variety, WH (or anything else) in position ① is incompatible with *that* in position ②; lifting this constraint (however this is done) allows *that*—but doesn't actually require it, since *that* is generally optional (or, as many would have it, in alternation with a zero complementizer), as in complement clauses and relative clauses. End of story.

Things are more complex than this, however. Even for complement and relative clauses in standard English, marking with *that* and lack of marking are not in 'free variation', as Zwicky (1994) notes, citing well-known data. Though the two structures are both available in most contexts, there are contexts in which only one or the other of them is. The picture that emerges is one in which a *that*-marked complement construction and an unmarked complement construction co-exist, ditto a *that* relative construction and a zero relative construction. The alternative constructions are semantically equivalent—so that it is no surprise that their distributions overlap so much—but have different virtues: the marked subordinate constructions provide a clear signal of clause organization for the hearer (the complementizer that marks the beginning of an embedded clause); their unmarked parallels serve the function of brevity for the speaker. Nevertheless, each is a construction on its own, with its own distribution in larger syntactic structures.

The same is true for doubly marked interrogative structures and singly marked ones in modern (non-standard) English. They are se-

mantically equivalent, so we should expect their distributions to be very similar. But in fact their distributions are different in various details, as we have seen: double marking is not available in non-finite clauses (Section 2.2) or in combination with subject-auxiliary inversion (Section 2.3).⁷ Again, the picture is of separate constructions, each with its own distribution in larger structures.

When we turn to languages other than modern English, once again (as with structure-specificity and the LHR) we find a failure of parallelism. Among the languages Radford mentions, Bavarian looks very similar to a variety of modern English in which the RZ examples are all acceptable; according to Bayer (1984:24): 'B[avarian] provides impressive examples of doubly-filled COMP. As a rule, any finite embedded clause may be introduced with two COMP-positions. This holds, of course, also for relative clauses. My examples indicate that the complementizers *daß...* [for complement clauses] and *wo...* [for relative clauses] may be missing'. That is, double marking appears to be simply 'optional', in free variation with single marking.

But Bayer's discussion of the neutral complementizers (filling position ②) continues with the proviso, 'but in fact there are many speakers who almost never leave them away'. Which is to say, for some speakers of Bavarian double marking is (nearly) obligatory, not optional at all.

Among the languages Radford cites, at the other end of the scale of optionality from Bavarian is Norwegian, as discussed by Taraldsen (1978). According to Taraldsen, interrogative clauses usually have only single marking, only a WH word. But indirect questions in which the WH element is the subject of its clause are obligatorily doubly marked:⁸

- **direct question, subject:**

Hvem ser mest svensk ut?

'Who looks most Swedish?'

—but **Hvem som ser mest svensk ut?*

'Who that looks most Swedish?'

- **indirect question, non-subject:**

Jeg lurer på hvem du liker best.

'I wonder who you like the best'

—but **Jeg lurer på hvem som du liker best.*

⁷It's possible that for some speakers who find doubly marked complementation natural, there are contexts in which double marking is obligatory (and single marking excluded). But there would be no way to detect this on the basis of the data attested so far.

⁸This is the descriptive generalization. In Taraldsen's analysis, all embedded interrogative clauses are (underlyingly) doubly marked, but the complementizer *som* is deleted from nearly all such clauses.

- **indirect question, subject of more deeply embedded clause:**

Jeg lurer på hvem du synes ser mest svensk ut.

‘I wonder who you think looks most Swedish’

—but **Jeg lurer på hvem som du synes ser mest svensk ut.*

- **indirect question, subject of its own clause:**

**Jeg lurer på hvem ser mest svensk ut.*

‘I wonder who looks most Swedish’

—but *Jeg lurer på hvem som ser mest svensk ut.*

Here, there is no optionality whatsoever. Double marking and single marking are (apparently) in complementary distribution.

My purpose here is not to survey the relationships between double marking and single marking of subordinate clauses in the world’s languages—though this certainly should be done, if it hasn’t been already—but to point out that the various analogues that Radford cites for the modern English situation are, once you look at the details, not notably analogous.

The most general view of the situation would be that in any given context in a language, doubly marked subordination might be obligatory (excluding single marking), optional (in alternation with it), or excluded (only singly marked subordination being available). These relationships might be the default for the language, so that double marking might be generally obligatory (as, perhaps, for some Bavarian speakers), optional (as for some modern English speakers), or excluded (as in Norwegian). But it could all be structure-specific. Or various things in between.

3.2 Incompatibilities

Convenient though it is as a point of reference, the subordination template in 3.1.1 is not a theory-neutral description of the relevant structures. In particular, it incorporates a number of assumptions about elements that are mutually incompatible, elements that (universally or parochially) cannot co-occur. I now survey these briefly.

Complementizer *that* and relativizer *that*

Neutral subordinators for complement clauses and relative clauses (not to mention adverbial subordinate clauses) compete for position ② in the subordination template and so are predicted to be incompatible with one another.

It might seem that this should follow from their semantics, but that isn't necessarily so. There are constructions that combine something of the semantics of relative clauses (in particular, picking out a referent having the characteristics described in the subordinate clause) and something of the semantics of (interrogative) complement clauses (in particular, using the answer to a question posed in the subordinate clause). This is the case for 'concessive free relatives' in English; the underlined clause in *Whatever you had in your hand exploded* can be roughly glossed as 'the thing that you had in your hand [relative clause semantics], whatever that was [interrogative semantics]'. We might then expect that each component of meaning could be signaled by an explicit complementizer. This is not the way things work in English, standard or otherwise, or in any other language that I know of, but it would not seem to be excluded by semantic considerations alone.

In Bayer's description, Bavarian provides clearly distinct subordinators for complement clauses (*daß*) and relative clauses (*wo*). In principle, a language could have a considerable inventory of such subordinators, marking various types of complement clauses (say, complements to nouns, as in *the idea that pigs can't fly*, vs. arguments of verbs, as in *I know that pigs can't fly*), various types of relative clauses, and/or various types of adverbial subordinate clauses. And, of course, a language could have distinct subordinators for finite vs. non-finite clauses.

English goes about as far as you can get in the opposite direction, using *that* for subordinate finite clauses in general, in *that* complements and *that* relatives. Whether this is, as many analysts assume, a single element with multiple uses, or two categorially distinct but homophonous elements, it seems quite clear that the relative *that* (unlike the relative WH words *which* and *who*) is not a pronoun of any sort.

Interrogative WH and relative WH

The subordination template puts XP[... WH], whether interrogative or relative, in position ①, and so predicts that they are incompatible with one another.

Again, this might seem to follow from the semantics, given that interrogative WH words are indefinite pro-forms, while relative WH words are definite pro-forms. But it's not difficult to imagine a function for doubly WH-marked clauses, like the underlined one in the invented *-Which people who came to the party amazed me*: one WH (*which* in *which people*) would be interrogative, and the other (*who*) would be anaphoric to the XP containing the first, reinforcing it and clearly marking the construction as subordinate, just as subordinators like

that do in WH+*that* constructions.⁹

Subordinating conjunction and WH

The subordination template also puts subordinating conjunctions like *if*, *because*, or *before* and XP[... WH], of any type, in position ①, and so predicts that they are incompatible with one another.

In this case, the interpretations I can imagine for such combinations are better conveyed by much simpler constructions, or else they are of extremely limited utility. Invented examples with indefinite WH in the scope of the subordinating conjunction, like **if who was at the party* 'if someone, whoever that was, was at the party' or **before which candidates I interviewed* 'before I interviewed some candidates, whichever ones that they were', would seem to convey nothing more than the much less complex *if anyone was at the party* and *before I interviewed any candidates*. And invented examples with the subordinating conjunction in the scope of definite WH, like **who if was at the party* 'the people who were at the party, if they were at the party' or **which candidates before I interviewed* 'the candidates which I interviewed, before I interviewed them', have interpretations that, insofar as they are coherent at all, would not be of much use.

I conclude that, on the basis of their semantics, these combinations of clause-introducing elements would be, at best, extremely infrequent in the languages of the world.

Inverted auxiliary and *that*

One further prediction of the subordination template, for those analysts (Radford among them) who have inverted auxiliary verbs located in position ②, is that inverted auxiliaries are incompatible with *that*. Having *that* and an inverted auxiliary fill the same position provides a direct account of why WH+*that* clauses are incompatible with auxiliary inversion.

Now, this part of the subordination template is not very solid as a universal. Inversion of subject and auxiliary verb (as opposed to subject and head verb, whether or not this verb is an auxiliary) is not common in the languages of the world, so that the incompatibility of *that* and

⁹There is at least one other way for an interrogative WH word to occur with a relative WH word, in examples like *Which candidates who came from France did you prefer?* But these are not instances of the subordination template; instead, the entire relative clause is a postnominal modifier, parallel to *from France* in *Which candidates from France did you prefer?* or *educated in France* in *Which candidates educated in France did you prefer?*

an auxiliary could perfectly well be particular to English. In addition, verb inversion is so frequently associated with the main-clause phenomena of fronting of topicalized or focused elements that it would be no surprise to find inversion excluded from subordinate clauses, such as those marked by *that* in English.

Confining ourselves to English, the fact that subject-auxiliary inversion (SAI) doesn't co-occur with *that*-marked subordination is remarkable only in frameworks where elements move (or are located) freely, subject only to general formal conditions—frameworks in which SAI and no inversion (with the subject before its predicate VP (SVP), instead of having its predicate VP wrapped around it as in SAI) are equal alternatives. If, instead, we look at the details of SAI, vs. SVP, in English, we see that SVP is overwhelmingly the default for clauses in general, SAI being confined to a small heterogeneous collection of constructions, some main-clause (such as yes-no questions, direct WH questions, and clauses with certain fronted focused elements, as in *Not a bite would I eat* and *From such humble origins did these remarkable consequences arise*), some subordinate (such as tag modifiers like *so/as would Sandy* and counterfactual conditional clauses like *were I ruler of the world*).

Which means that incompatibility with SAI is no big thing in English. All clauses have SVP unless some construction stipulates SAI instead (or as an alternative). English WH+*that* clauses have the default subject-predicate rule, SVP, and only SVP, a circumstance that requires no special statement whatsoever.

What does require a special statement is which WH complement clauses have alternatives with *that*. Setting aside exclamations like (6), the inventory of WH complement constructions that have WH+*that* alternatives makes a natural class: finite subordinate interrogative constructions.

Note that despite the label 'concessive free relatives' (for clauses like the underlined one in *Whatever we had in our hands sparkled* 'Whatever it was that we had in our hands, it sparkled'), such expressions have the syntactic properties of interrogative constructions rather than relative constructions, free or bound: for instance, they allow expletive postmodifiers of the WH word (*Whatever the hell we had in our hands sparkled*; **What the hell we had in our hands sparkled*; **The stuff which the hell we had in our hands sparkled*). Concessive free relatives then count as finite subordinate interrogative constructions for the purpose of syntactic generalizations about English, including the generalization above about the WH constructions allowing WH+*that* alternatives.

In any case, the ungrammaticality of clauses like –*Which person*

that spoke first?, even in non-standard varieties, provides no support for a double-movement analysis of direct subject questions like *Which person spoke first?* For speakers of modern English in general, direct WH questions are incompatible with that marking. So far as I can tell, this is a contingent fact, just one of those things; it could have been otherwise.

Summary

Whether it's to be understood as a list of places where clause-introducing elements can be located, or as a list of landing sites for moved constituents, or as a static constraint on subordinate structures, the subordination template is intended as an entirely formal condition on syntactic structure. It has nothing to do with meaning.

I've now argued that, with the possible exception of the incompatibility of fronted interrogative WH and fronted relative WH, the subordination template has no justification in semantic terms. If there is a constraint here—universal or parochial—it's not because the elements involved are semantically incompatible with one another.

My guess about these matters is that the combinations barred by the subordination template are merely relatively unlikely in the languages of the world. Our failure to find counterexamples to it is, on this account, just the result of our not having looked very hard for them, or our not having collected enough data from a sufficiently wide variety of languages and dialects.

(On the other hand, if the subordination template turned out to be well-supported, it would be a stunning demonstration that there can be purely formal syntactic universals, not motivated by semantics or by discourse function. My money is against being stunned, but I could be wrong.)

4 History

Ok, so there's this phenomenon for some speakers of modern English. How could it have come about? Why would there be an option for *that* marking in certain finite (and largely subordinate) interrogative WH constructions, in particular those in which the WH word is a modifier of some lexical head?

4.1 Initiation

The story I'd like to tell has two possible, and not incompatible, starting points, one having to do with perception, the other with production. Both stories take off from subordinate clauses like *what kind of people that they are dealing with* (12), as an alternative to *what kind of people they are dealing with*, and *how much water that you need to drink* (13), as an alternative to *how much water you need to drink*.

(a) The perceptual story is that someone hearing *how much water you need to drink* might (mis)parse it as involving a head N *water* with a postnominal zero relative *you need to drink* __ (rather than treating *water* as fronted material within a complement clause). In which case the hearer might want to supply an explicit marker of relativization, *that*, for lexically headed subordinate WH interrogatives. Yes, *which* would also be possible, but *that* is the stylistically unmarked alternative.

On the side of the perceptual story is the fact that most of the examples in the RZ corpus have 'hidden question' parallels involving definite articles (rather than WH determiners) in combination with following relative clauses, for which the explicit marker *that* is available. Here are a few of the parallels:

- (1) I'm not sure what kind of ban that FIFA has in mind.
- (1') I'm not sure of the kind of ban (that) FIFA has in mind.
- (2) We'll see what sort of pace that Daley Thompson's running at.
- (2') We'll see the sort of pace (that) Daley Thompson's running at.
- (16) What this is about is providing them with whatever assistance that they need.
- (16') What this is about is providing them with the assistance (that) they need.
- (15) It all depends on how much work that you had to do.
- (15') It all depends on the amount of work (that) you had to do.
- (19) ... until late in the week, when we see how many people that were arrested ...
- (19') ... until late in the week, when we how the number of people (that) were arrested ...

Such parallels are not available for every one of the original examples—*However many people that were there, ...* (8) and *Look at it as a tribute to how deeply that I feel about you* (17), for instance, lack

close parallels—but they are sufficiently frequent to serve as models for a parsing of embedded interrogative WH constructions as involving a relative clause.

(b) The production story is the one I sketched in Section 2.3 above: the speaker begins a long fronted NP[... WH], to be followed by a complement S/NP, but on reaching a lexical noun like *water* (which could serve as the head for a relative S/NP), switches to a postnominal relative construction and supplies the stylistically unmarked subordinator *that*. The speaker then produces a syntactic blend, with two distinct constructions overlapping one another, as in this example from the CBC radio program ‘As It Happens’ on 8/31/00: *While I don’t condone the actions of the Canadian government, but I think that...*

In either story, we are considering possible triggering events. In the perceptual story, the original hearer becomes a source (in productions after the misparsing) of a certain number of WH+*that* clauses. In the production story, the original speaker is the source of one such clause. What has to happen now is that there are enough occurrences of these clauses for some hearers to take them to be simply part of English.

In either story, the examples presented to new hearers will have the properties listed in Section 2 above. They will involve interrogative-type, rather than relative-type, constructions (Section 2.1); WH relatives will already have a marker of subordination, namely the WH pronoun itself. They will involve finite constructions (Section 2.2), since only finite clauses have alternatives with initial *that*. They will be mostly subordinate, and certainly not inverted (Section 2.3): mostly subordinate, since that’s where the alternation between zero marking and *that* marking otherwise occurs in English; and not inverted, since SAI is a constrained alternative to the default SVP in English, and doesn’t otherwise occur in *that*-marked subordinate clauses.

The WH word in these examples will be a modifier to a lexical head word, and that word will be a noun, so that the fronted XP[... WH] will have at least two words in it (Section 2.4). This follows in either initiation story, because a lexical noun is required as a bridge between an interrogative-type construction and a restrictive relative construction.

The occurrence of *that* will be ‘optional’ (Section 2.5), because hearers will be supplied with plenty of instances of ordinary, *that*-less, interrogative constructions, from other speakers and even from the initiating speakers. And there will be no restriction on the syntactic function of the fronted NP within its subordinate clause (Section 2.6), since there

is no such restriction for *that* relatives; in particular, there are subject *that* relatives, as in *the people that were there*.

4.2 Another source?

There is at least one other syntactic structure in English that might have served as a model for WH+*that* clauses. This is the NP[... WH] plus postnominal *that* relative (as in the underlined portion of *Which people that you went to college with were at the party?*) that I've mentioned a couple of times already.

This is not a very likely model, for at least two reasons. First, the structure is entirely natural with a single-word head: *Who that you went to college with was at the party?* Second, it is entirely natural with a WH relative: *Which people who you went to college with were at the party?* Neither of these facts is mirrored in the WH+*that* clauses; the attested examples include none with a single-word WH expression and none with WH+WH (*-I don't know how much water which you need to drink*).

4.3 The new grammar

I now ask what grammar people might construct on the basis of examples provided in the initiation phase.

The amalgam analysis

One possibility is that this grammar is a direct reflection of the perceptual or production errors that gave rise to these examples, that it licenses expressions that are amalgams of an interrogative construction with a relative construction, so that (like English verbal gerunds as analyzed by Wescoat 1994 and Malouf 2000, in rather different conceptual frameworks) they must satisfy the conditions on both of the contributing constructions.

On this account, in WH+*that* clauses, the XP[... WH] in position ① and the constituent Z (of category S/XP, that is, S missing an XP) together must make an interrogative-type WH construction, while at the same time, the filler in position ② and the constituent Z together must make a restrictive *that* relative serving as a postmodifier of the head word in the XP[... WH]. The semantics for this construction is that of the interrogative-type WH construction; the relative construction contributes only its formal requirements.

It might seem that this amalgam analysis predicts, rather than stipulates, the LHR, but that is not the case. As I've observed several times,

single-word WH expressions can have postmodifying relative clauses, so that the analysis as it stands would allow unattested types of examples, like *-I don't know who that I should ask*. I conclude that the LHR must be stipulated: the filler in position ② and the constituent z together must make a restrictive *that*-relative serving as a postmodifier of the lexical word in the XP[... WH] that is modified by the WH word there.

Notice that the amalgam analysis also stipulates, rather than predicts, that the relative clause construction is a *that*-relative. I grant that the examples that serve as the basis for positing a new construction all have *that* rather than a relative WH word. But hearers might well conclude from this evidence that all that's required is some explicitly marked relative construction, in which case they should themselves produce some examples like *-I don't know which people who I should ask* as well as some like the attested *I don't know which people that I should ask*. Why should all the hearers be so conservative?

A final problem with the amalgam analysis is that it fails to extend to APs as well as NPs (and PPs). Since adverbs like *soon* do not serve as heads of restrictive relative clauses, the amalgam analysis does not predict the possibility of examples like (21) *I am ... surprised at how soon ... that the judge approved it*. There is some suggestion that there are speakers for whom WH+*that* clauses are indeed restricted to NP contexts, but I myself now find examples like (21) so ordinary that I fear that I'm failing to notice them.

I wouldn't want to claim that amalgam-style analyses are out of the question. In fact, I believe they are called for in other circumstances. But certainly they are complex, and it seems reasonable to suppose that people wouldn't posit such analyses unless there was considerable evidence in their favor.

A special *that*-marked construction

So the question for WH+*that* clauses is whether there is a less complex, and at least equally well supported, analysis.

I believe there is. Such clauses can be seen as *that*-marked alternatives to ordinary interrogative-type WH clauses; the alternation is quite parallel to that between *that*-marked and unmarked complement clauses (rather than to that between *that* relatives and zero relatives). This analysis has some of the flavor of Radford's, but it is construction-specific in an essential way.

A crucial assumption is that a single construction can serve as an ingredient in a number of different larger constructions, each of which could place further idiosyncratic stipulations on its ingredients. An-

other crucial assumption is that constructions with the same semantics will, as a default, serve as alternatives to one another in such a composition of constructions. The idea is then that there are (at least) two alternative constructions that can combine with a preceding XP[... WH] to yield an interrogative-type clause: one of the form S/XP, one of the form *that* + S/XP. The latter combination—XP[... WH] + [*that* + S/XP]—is idiosyncratically subject to the LHR.

This is a bald-faced stipulation. It is a mark the construction bears of its historical origin. The construction might eventually be generalized, but for the moment you can see some signs of its historical origin.

In any case, the analysis I've just sketched is no more stipulative than the amalgam analysis, and it is entirely consistent with AP rather than NP as the fronted XP[... WH].

The (extremely hypothetical) account I've given here turns on relative clause constructions as the source of the WH+*that* phenomenon but posits complement constructions as the crux of its analysis.

5 A wider sample

My discussion has been based on the slimmest of empirical supports: just 29 attested examples; plus judgments on a small set of examples from one speaker who doesn't reject WH+*that* clauses in general, and some comparative judgments (distinguishing things-I-wouldn't-say from things-I-can't-believe-anyone-would-say) from a number of other speakers, including myself; to which I have added a number of hypotheses about which gaps in the evidence are accidental and which probative, and about how the data might extend to new contexts.

This won't do. So much of it could be wrong, in details small and large. What we need is a much larger bank of attestations, and (given that examples are likely to be very infrequent) collections of judgments from speakers for whom at least some types of WH+*that* clauses are natural. Fortuitous collection of examples is still worth doing, but it's probably not going to net enough new examples.

Fortunately, some progress in exactly this direction has now been made by Seppänen & Trotta (2000), with their 90 WH+*that* examples from truly gigantic corpora of (mostly British) English. There are fresh problems that arise from looking at very large corpora; in particular, it's virtually a sure thing that some of the examples are just production errors, and it won't always be easy to tell which ones. Informant judgments can help, but only if the informants are closely matched in variety to the original speakers/writers, and that's hard to ensure.

It is also unfortunate that Seppänen & Trotta (S&T) don't provide the actual examples, but only (for the most part) statistical summaries. Even without information about the context or about the people who were the sources of the examples, the sentences themselves would be valuable.

Nevertheless, there's a lot to be learned from their data, even if much further analysis is called for.

5.1 Interrogative, not relative

In Section 2.1, I formulated the descriptive generalization that WH+*that* occurs only in interrogative constructions; explicitly excluded are relative constructions (restrictive relatives, appositive relatives, and free relatives).

S&T's Table 7 (pg. 171) breaks down their WH+*that* examples by clause type. Only 27.8% of these are labeled as interrogative, an additional 6.7% as exclamative. Exclamatives share many more of their properties with interrogatives than with relatives, so it makes sense to lump them together, giving a total of 34.5%, a figure that is apparently overshadowed by the 53.5% of the examples that are labeled as 'free relatives'. However, all the examples they give that they identify as 'free relatives' are in fact concessive free relatives (with WH-*ever* words in them), which I would classify as interrogative (Section 3.2.4 above). If this carries through to their data in general, then all but 12.2% of their WH+*that* are interrogative.

The recalcitrant 12.2%, 11 examples, make a rather puzzling lot. Almost all of them, 9 of the 11, are also exceptions to the Lexical Head Restriction (Section 2.4; see 5.3 below), and S&T report that their informants treat all 11 of them as marginal at best. The two examples that satisfy the LHR are both (probably) appositive relatives, for what that's worth: *Philip Hayley, the main character of the novel, undertakes an excavation of the life of his charlatan father, around whose numerous exploits that the plot revolves; She did different things with us like doing dancing stuff like that which things that we hadn't done before* (pg. 172).

The question here is whether these relative-clause examples represent an extension, for some speakers, of the WH+*that* pattern to relative clauses, or whether they are merely noise in the data.

5.2 Embedded, not main

All of S&T's 90 examples can be assumed to be in embedded clauses, with the possible exception of the 'exclamative' category. My guess is

that these are main-clause exclamations, like the 'Irish' *What a mine of useless information that I am!* (RZ example (6)). If so, we need to determine whether they represent a exclamative construction characteristic of Irish English (just who were the speakers/writers of these sentences?), whether they are production errors, whether they are some entirely new phenomenon, or whether they include examples of several different types.

5.3 The Lexical Head Restriction

In Section 2.4 I noted that the RZ corpus had no single-word WH phrases (*I don't know who that did it*) in them, and that informants tended to reject such examples. On the whole, S&T's discussion reinforces the LHR, especially in the judgments of their informants. Still, some 27.8% of their 90 WH+*that* examples (25 items in their collection) have single-word WH phrases in them, and this makes for a certain uneasiness. S&T (171) suggest that their examples (... *this is the outfit who that could live more easily with its second string; ... a few of the trickier questions which that often regularly come my way; ... in challenging situations where that they've gotta rise to the occasion; If I recall er when that er the King Street car park was given to the town ... ; ... like some old school friends erm this lady who that makes very nice coffee; I don't know why that you go for a certain colour ...*) indicate "a highly restricted use as a feature of certain idiolects only".

It's not clear what such examples represent. A certain number of production errors. A certain amount of Irish (or other regional) English. Previously unappreciated details of the WH+*that* construction(s). An extension of the modern WH+*that* interrogative complement construction to contexts approximating the range of its occurrence in earlier forms of English. Or even the earlier construction maintained without essential change in some varieties, during a period when this construction fell out of use for most varieties.

The gross outlines of the modern construction do seem clear. It is (mostly) restricted to interrogative main-clause finite WH contexts where there is a lexical head. Its details do not follow, for the most part, from the semantics and pragmatics of its parts (though these are scarcely irrelevant). It occurs in both spoken and written English, in both colloquial and formal contexts (though perhaps not with the same effect in all these contexts). It's infrequent, for several good reasons, but it can't (always) be dismissed as a production error. And it deserves further attention.

References

- Bayer, Josef. 1984. Towards an explanation of certain that-t phenomena: The COMP-node in Bavarian. In W. de Geest and Y. Putseys, eds., *Sentential complementation*, pp. 23–32. Foris, Dordrecht.
- Henry, Alison. 1995. *Belfast English and standard English*. Oxford University Press, Oxford.
- Malouf, Robert. 2000. Verbal gerunds as mixed categories in Head-driven Phrase Structure Grammar. In Robert D. Borsley, ed., *The nature and function of syntactic categories*, volume 32 of *Syntax and Semantics*, pp. 133–66. Academic Press, San Diego, CA.
- Radford, Andrew. 1988. *Transformational Grammar*. Cambridge University Press, Cambridge.
- Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics*, 33(2):431–83.
- Seppänen, Aimo. 1994. Subject gap phenomena in English and Swedish. In C. Hedlund and A. Holmberg, eds., *Proceedings of the XIVth Scandinavian Conference of Linguistics and the VIIIth Conference of Nordic and General Linguistics*, pp. 127–44. Department of Linguistics, Göteborg.
- Seppänen, Aimo and Joe Trotta. 2000. The *wh- + that* pattern in present-day English. In John M. Kirk, ed., *Corpora galore: Analyses and techniques in describing English*, pp. 161–75. Rodopi, Amsterdam.
- Taraldsen, Knut Tarald. 1978. The scope of *wh* Movement in Norwegian. *Linguistic Inquiry*, 9(4):623–40.
- Westcoat, Michael. 1994. Phrase structure, lexical sharing, partial ordering, and the English gerund. In *Proceedings of the Berkeley Linguistics Society Annual Meeting*, volume 20, pp. 587–98.
- Zwicky, Arnold M. 1986. WH constructions in English. In *Ohio State University Working Papers in Linguistics*, volume 32, pp. 114–24.
- Zwicky, Arnold M. 1994. Dealing out meaning: Fundamentals of syntactic constructions. In *Proceedings of the Berkeley Linguistics Society Annual Meeting*, volume 20, pp. 611–25.