

SOME FACTS ABOUT *WHO* AND *WHOM*

Fred W. Householder

In this paper we consider the relative frequencies of the four basic types of human objective relative clauses in English (The woman I love, The woman that I love, The woman whom I love, The woman who I love) in a variety of literary sources for colloquial speech from Shakespeare to the present, and attempt to relate these to Klima's four types or stages of English (L_1 , L_2 , L_3 , L_4). Evidence is offered to show that Klima overlooked an important type with *I*, *he*, *we* in the places where L_4 has *me*, *him*, *us*, regardless of government or agreement, here called $L_{2.1/2}$. In all or nearly all sources the clause-type represented by "The woman who I love" is the rarest and either "The woman I love", or (in George V. Higgins and Mark Twain) "The woman that I love" is the most frequent type. The type (mentioned by Klima in a footnote) which regularly employs "He asked who I loved" but prefers "The woman whom I loved" to "The woman who I loved" (here called $L_{1.1/2}$) is shown to be clearly preferred to Klima's L_2 from Shakespeare on down to the present, as well as to L_1 .

A more complete title would be "Some facts about *who*, *whom* and other relatives referring to humans, and some speculations about lexical variation and linguistic change, as well as notes on surface "case" in all pronouns." All of these topics I have at least touched upon before. In Householder 1969 (a review of a book by R. W. Langacker) I discussed the uses of *who*, *whom*, *that* and zero as object relatives and prepositional object relatives, and in Householder 1971, pp. 5-7, I discussed the competing rules in spoken English for the choice between *I* and *me*, *she* and *her*, *we* and *us*, etc.

Earlier discussions of some of these questions which will be referred to occasionally are those by Jespersen (1940, 1949_a, 1949_b); Sapir (1939: 166-179) (he sees direction in the variation, "Drift"); Klima (1964) whose table 1 (below) must be contrasted with the table 1 reported from Klima's 1964 thesis (which I have not seen) by Lightfoot 22; Bever and Langendoen 33-8, 49-76.

What are the problems with these treatments that lead me to look at the situation again? There are, in fact, several dubious assumptions, anecdotal proofs, confident assertions and questionable conclusions, not just in these works, but in a great deal of the treatment of English relative clauses of the past fifty years or so. What first caught my eye was the uniform assumption in many publications of MIT linguists that the normal, neutral or basic restrictive personal object relative was "who" -- "The woman who Peter loved", "the woman who Peter talked to". I think we will be able to show that that is, in fact, one of the least frequent forms and also one of the most highly

colored (though mere statistics cannot show this), carrying connotations of pretentiousness, artificiality or careless improvisation. The choice of the *wh*-form signals literariness or elegance, but the choice of *who* over *whom* simultaneously signals almost the opposite; it's a schizophrenic type, (in non-restrictive oral use, it's much closer to normal, since, for most people, neither *that* nor zero is possible in that environment.)

Let's list a few assumptions to be tested. (1) The interpretation of Klima's table 1 :

Klima's Table 1

	L ₁	L ₂	L ₃	L ₄
A 3.01	She could see him near me.			
3.02	He and I left.			Him and me left.
3.03	We two left.			Us two left.
3.04	We all left.			
3.05	It was I.		It was me.	
B 3.06	Could she see him?			
3.07	Who could see him?			
3.08	Whom could she see ?	Who could she see?		
3.09	With whom did he speak?			[—]
3.10	Whom did he speak with ?	Who did he speak with ?		
3.11	Who was it ?			
3.12	Who was the leader ?			
C 3.13	He knew whether she could see him, or not.			[—]
3.14	He knew who it was.			
3.15	He knew who was the leader.			
3.16	He knew whom he spoke with.	He knew who he spoke with.		
3.17	He knew with w h o m h e spoke.			[—]

- D 3.18 The leader who
could see him
left.
- 3.19 The leader The leader who I
whom I saw saw left.
left.
- 3.20 The leader with [—]
whom he
spoke left
- 3.21 The leader The leader who
whom he he spoke with
spoke with left.
left.

reported by Lightfoot p. 22 labels L_1 as "late Middle English", L_2 "18th century", L_3 "late 18th century colloquial", and L_4 "a dialect form of Modern English"; as given in the *Language* article, Klima 1964 (p.3, note 3), L_1 is "general school grammar for formal writing", L_4 is "the dialogue in Nelson Algren" (1956 and 1947), whereas L_2 and L_3 are intermediate styles, each being a little less formal than the one before. Klima also remarks that there is another style between L_1 and L_2 in which "the interrogative pronoun does not have case marking, except after prepositions, whereas the relative pronoun has case marking under the same conditions as L_1 ."

Let us consider the historical interpretation first. According to it, both relative and personal pronouns receive case-marking uniformly according to what I will call the Latin rule. This is probably true for the personal pronouns, with the exceptions mentioned by Mustanoja, 113-114, 122, 124-5, 129-130, 133, 134, some of which are like Klima's L_4 (with *me* for *I*), others like a competing dialect not mentioned by Klima, in which *I* occurs for L_1 *me* (e.g. *between you and I*-found in Shakespeare's *Merchant of Venice* 3.2.320). But the problem with *who* and *whom* in Middle English is that as relatives they do not become current till the end of the period, and even then *who* (nominative or not) hardly occurs at all--never in Chaucer, according to Jespersen (1949a) 121 (cf. Mustanoja 199-200), though *whom* does, most often immediately following a preposition (see Mustanoja 201, Steinki (1932) 34-5; note also the AV(King James Bible) rule in Jespersen (1949a) 118). When *who* does occur in late ME, it is mainly in two uses which are hardly known today, (1) as a connecting relative in the formulaic expression "...God, who have you in his keeping" i.e. "and may he have you etc."(see Steinki 40-43--note on 42 that three times people write "God, whom have you in his keeping") and (2) in indefinite-general relatives, either "headless" or with "antecedents" after the relative clause, e.g.(from Caxton) "who pretendeth to God, God attendeth to him" (see Steinki 45-6). Something like both of these uses is quite common in Shakespeare (at least 129 connecting relatives and 84 general)--both put *who* initial in sentence or (co-ordinate) clause, where no other relative word can

stand, as a rule. At any rate, parallels for the relative clauses in L_1 (except D 3.20) are virtually non-existent in ME. The questions and relatives of L_2 , on the other hand, can easily be paralleled in Shakespeare (who also sometimes uses the types in L_1). But Shakespeare speaks the dialect mentioned in Klima's note 3 (also noted by Jespersen and Sapir, who [166 n. 10] does not mention the alternative object form *who*) which differentiates case-marking in interrogatives from case-marking in relatives¹: Shakespeare's interrogative human object pronoun is *who* 42 times, *whom* 40 (if we exclude the position immediately after a preposition, it is *who* 34 times, *whom* 11), whereas the relative is *who* 19 times and *whom* 230 times. This is a clear and substantial difference. After a preposition the relative is always *whom* (81 times), but the interrogative shows *who* 29 to *who* 8. For Shakespeare, as for us (and for ME as well) with prepositions immediately preceding the relative, *whom* is categorical. For us (unlike Shakespeare) interrogative *whom* is only a little less categorical in that environment. Klima's L_3 differs from L_2 only in putting *me*, *him*, etc. after the verb "be" ("it's me" etc.); Mustanoja 133 cites "I him am" and "art thou him" from the 14th century and Shakespeare has one example that I can find (*Lear* 1.4.184 "I would not be thee" spoken by the Fool), so in this case Klima's date may be OK. Certainly L_4 ("Him and me did it, us two did it") has no examples at all in Shakespeare, but instead "between you and I" and "the hand of she here" (*Antony and Cleopatra* 3.13.98). I used to think these nominatives were hyperurbanisms, overcorrecting for the accusatives of L_4 ; but this variety (call it $L_{2,1/2}$) is earlier than L_4 , if we may judge by Shakespeare and his contemporaries.

There are between 5 and 7 (depending on how you count them) special environments (disjunctive, to use the French term) in which the Latin rule (of L_1) may be overridden to yield either $L_{2,1/2}$ or L_4 :

1. The position as predicate with "be" (and possibly some other such verbs) which Klima uses to define L_3 is different from the rest; we do not look at one who says "it's me" in the same way as if he said "me and him are friends", or "it's I" as if he said "between you and I." Shakespeare here almost always uses "I"--as he does in the other disjunctive positions--which, in this case, agrees with L_1 , the Latin rule. A real $L_{2,1/2}$ here would be "I want it to be I."
2. The position in conjoined NP's, linked by (a) *and*, *as well as*, *or* (e.g. "between you and I") or (b) *than*, *as*, or *but* (after negatives). In this situation Shakespeare has *I* against the rule 6 or 7 times, but never *me* against the rule. L_4 does not show itself here (in the *a* environment) in Shakespeare, though it does at least once (*Antony and Cleopatra* 3.3.14

¹ The OED agrees with Sapir and Jespersen: under WHO 15, the interrogative, they say "common in colloquial use as obj. of a verb, or of a preposition following at the end", at II, 13, the relative. "now rare or obs."

- "Is she as tall as me?") in the *b* environment. Elsewhere both *a* (7 times) and *b* (20 times) are $L_{2\frac{1}{2}}$ in S. (e.g. *As You Like It* 1.1.166), though the relative is only *than whom*, in *Love's Labours Lost* 3.1.78.
3. The situation in which a pronoun has an adjective before it or a noun phrase or modifier after it (*Poor me! Us boys, them there*) generally conforms to L_1 in Shakespeare, except for a few $L_{2\frac{1}{2}}$ examples like "The hand of she here" in *Antony and Cleopatra* 3.13.98, and "to poor we thine enmity's most capital" *Coriolanus* 5.3.103.
 4. Where the pronoun immediately precedes the relative (as antecedent), there are two subtypes, one (a) closely resembling type 3, where the modifier is a relative clause, (once L_4 , in *Macbeth* 5.8.34 "damn'd be him that first cries 'hold, enough!' ") and the other (b) where the whole NP undergoes emphatic fronting (always $L_{2\frac{1}{2}}$ e.g. *Timon* 4.3.39-40 "She whom the spital-house...would cast a gorge at, this embalms"). Strangely enough, though nowadays we seldom use personal pronouns as antecedents, when we do we may slip into $L_{2\frac{1}{2}}$. A recent (April 1986) column (*The Writer's Art*) by James Kilpatrick lists a few "Horrid Examples", most of which are $L_{2\frac{1}{2}}$ of type 2a, and one type 4 "There is a fine array for she who craves grown-up innocence". Since the relative which follows is generally a subject, this can often be considered inverse attraction (antecedent attracted to case of relative), as is true for all twelve examples in Shakespeare. However, in today's usage I would not be surprised to see a sentence like "This is for she whom we all admire." Kilpatrick's other $L_{2\frac{1}{2}}$ cases include 5 of type 2, 4 joined by *and*, one by *or*, and one of type 3 ("of we young bucks").
 5. One more type which occurs in Shakespeare ($L_{2\frac{1}{2}}$) and in ME (see Mustanoja 113, "for us, we sinful folk"), though it is probably rare today (I cannot cite an example) is the occurrence of a personal pronoun in apposition to a noun phrase. Exclamatory isolation is probably a subtype of this (the isolated pronoun is somehow appositional or a repetition), and occurs also today ("who's to take care of him if not me?" in *Man with the Golden Arm* 94 is isolated by deletion). Compare, in *Henry IV Part I*, 1.2.133 "Who, I rob? I a thief?" Clearly Shakespeare's language is $L_{2\frac{1}{2}}$, while Algren's is (as Klima asserts) L_4 on this point. Compare also "Me, I don't like it" in common use.

We can summarize this briefly: for all varieties of English, from ME on, unstressed pronoun subjects are *I, thou, he, she, we, they*, while unstressed objects of verbs and prepositions are *me, thee, him, her, us, them*, provided they are neither conjoined nor modified. In all other situations (those identified by numbers 1 to 5 above, and perhaps a few others), Shakespeare mainly has $L_{2\frac{1}{2}}$ (i.e. *I, thou, he*, etc. everywhere), modern literary usage has L_3 (after *be*) and L_4 in situation 2b (after *as, than* or *but*), low class modern has L_4 throughout, and some "elegant" styles have $L_{2\frac{1}{2}}$ again, either throughout, or at

least in environments 2a and 4a. The choice between *who* and *whom* resembles this situation only for objects of immediately preceding prepositions, where all styles have L_1 (i.e. *whom* like *me*, etc.) though interrogatives (never relatives, either in Shakespeare or today) may sometimes shift to L_2 (*with who?*) especially when nothing follows the pronoun. No one rule will possibly work for all three pronouns in any dialect from Shakespeare on.

Both $L_{2\frac{1}{2}}$ and L_4 can be defined without mentioning syntax or semantics: $L_{2\frac{1}{2}}$ uses *me* etc. for unstressed, unconjoined pronouns immediately after verbs or prepositions, and elsewhere *I* etc. L_4 uses *I* etc. for unstressed unconjoined pronouns immediately before verbs or auxiliaries and immediately after auxiliaries, and elsewhere *me* etc. Few people speak a pure form of either dialect.

So, if we consider the order L_1, L_2, L_3, L_4 to be chronological, it can only be so for L_3 and L_4 (and not provably even for them); L_1, L_2 and $L_{2\frac{1}{2}}$ are all simultaneous, coexisting happily in Shakespeare, though the relative types of L_2 at the bottom (3.19 and 3.21) are rare in Shakespeare.

(2) If we reinterpret Klima's table 1 stylistically (as in *Language*), our luck is better. Almost every grammarian since 1900, however, including Fowler and Fowler, allows L_3 (*me* after *be*) to be acceptable, and *than whom*, which is not obviously placed anywhere (marginally L_1 , perhaps), is universally used, following Shakespeare (*Love's Labours Lost* 3.1.178) *who* never has *than him* for *than he*. On the other hand, in modern usage *me, him* etc. in disjunctive environments do not always signal low class Chicago bums, particularly in environment 2b (I have yet to find a native speaker who will accept "He came to the costume party disguised as I") and 3 exclamatory (*Poor me!* beats out Shakespeare's *Poor I!*); and $L_{2\frac{1}{2}}$, ignored by Klima, also carries some stigma, at least for many of us, like Kilpatrick.

Suppose we turn again to the claim (e.g. Klima's L_2) that the choice between *who* and *whom* is made by the same rules, whether we have a relative or an interrogative use. I have already remarked that this is not true of Shakespeare, but neither is it true in a wide range of modern corpora, some counted by others, some by me.²

² The principal published count is in Quirk's "Relative clauses in educated spoken English", in Quirk 1968 94-108, (Table on 104), reprinted from *English Studies* 38(1957) 97ff.

There is a partial count of a minuscule sample (82 relative clauses of a juvenile dialect spoken in Reading) in Cheshire 1982, 72-75, (the only true spoken corpus besides Quirk's) but the raw data are not given, and it is difficult to recover them from the frequency indices of table 33. Some information about Middle Scots is also available in Romaine 1982, as well as more on the history of English relative. My own counts were made from the plays of Shakespeare (counting directly and also checking with the published lexica and concordances of Spevack, Schmidt and Bartlett), the Brown University millionword corpus (using a KWIK-type concordance; no figures are available for the zero-relatives, -of course), a small sample from the AV (*Numbers* 16-19, *Ezekiel* 16-19, and *Acts* 16-19), and a long list

Here is a table to show the numbers of *who* and *whom* (object of verb or of dangling preposition) questions and relative clauses in these various texts. Objects of immediately preceding prepositions outside of Shakespeare are invariably *whom* for both relative and interrogative, except for a single example in Dick Francis (*Whip Hand* P.9 "For who else would I risk coming here?") Clearly *who else* is a unit and cannot be case-marked to **whom else*.

It is clear that from Shakespeare on down, interrogative *who* is favored over *whom* at least five to one, except in literary prose (I am surprised that there are not some *who* questions in the Brown U. corpus, which does have a few conversational selections). On the other hand, it is only in the nonrestrictive use in British mysteries that relative *who* outpolls *whom* significantly, and in several texts *whom* is a substantial winner. Clearly the rules are different, and Klima's note 3 applies to all these texts except the Bible and the Brown

of fiction, mainly detective fiction as listed below. Algren is included because Klima used him to define L₄, Linington and Higgins because they use large amounts of naturalistic dialog, though not all of it is L₄; and they do (like Algren) use zero subject relatives (e.g. "There's a man downstairs wants to see you" and the like (as mentioned in Bever and Langendoen 1969 note 11). Both Higgins and Linington make some use of style to characterize individual speakers (more than Algren)--a device used much less frequently by Shakespeare, whose main trick is to give his comic low-class characters lots of malapropisms. The first nine entries on this list give a nearly equal sampling of U.S. and British mystery writers. (In addition, I have a number of *who* and *whom* object clauses collected from other mysteries at an earlier date, but without detailed documentation or counts of any other sort of relative clause. The figures for these in later tables are labeled *Old Notes*.)

The next four are Linington novels (three published under the pen name Dell Shannon), the next three Higgins novels (none of these are mysteries, though some of his earlier books are crime novels), and the last two Algren.

CORPORA

1. *Ellery Queen's Mystery Magazine* 1985 July, Aug., Sept., Oct., Nov.
 2. *Alfred Hitchcock's Mystery Mag.* 1985 June, Aug., Sept., Oct.
 3. *Ellery Queen's Prime Crimes* 3, Anthology #53, Fall 1985
 4. *Espionage Magazine* 1985 August & November
 5. Dick Francis *Whip Hand* 1981 Pocket Books N.Y. (Br.)
 6. Dick Francis *The Danger* 1985 Fawcett Crest, N.Y. (Br.)
 7. Simon Brett *Situation Tragedy* (A Murder Ink Mystery) 1981 (Br.) A Dell Book, N. Y.
 8. Andrew Garve *Counterstroke* 1979 Penguin Books, N.Y. (Br)
 9. Julian Symons *The Name of Annabel Lee* 1984 Penguin Bks., NY (Br)
 10. Elizabeth Linington *Greenmask* 1964 Popular Library, NY (US)
 11. Dell Shannon(=E.L.) *Death of a Busybody* 1985 (1963) Mysterious Press, NY (US)
 12. Dell Shannon(=E.L.) *Knave of Hearts* 1984 (1962) Mysterious Press, NY (US)
 13. Dell Shannon(=E.L.) *Double Bluff* 1985 (1963) Mysterious Press, NY (US)
 14. George V. Higgins *A Year or so with Edgar* 1980 Berkley Books NY
 15. George V. Higgins *A City on a Hill* 1985 (1975) Carroll and Graf NY
 16. George V. Higgins *A Choice of Enemies* 1985 (1983) Carroll and Graf NY
 17. Nelson Algren *The Man with the Golden Arm* 1956 (1949) Fawcett Crest NY
 18. Nelson Algren *The Neon Wilderness* 1947 (1933) Doubleday & Co. NY
- [The dates in parentheses are dates of first publication]

Table 1

	Question		Nonrestrictive Relative		Restrictive Relative	
	who	whom	who	whom	who	whom
Shakespeare	34	11	16	144	19	230
Bible (AV)	0	0	0	8	0	18
Old Notes US	[no data]		2	0	9	8
Old Notes Brit.	[no data]		9	1	21	15
US Mysteries	14	2	3	2	10	14
Br. Mysteries	13	2	12	6	19	20
Brown U.	0	3	1	20	1	23
Quirk	[no data]		1	1	4	9
Algren (dialog)	14	1	0	0	0	1
Higgins	26	1	1	4	3	12
Linington	9	0	0	7	3	0

University Corpus, which seem to be pure L_1 -nothing is L_2 at all. Incidentally, no support is offered to those who are reported to believe that the British prefer *whom*, which good Americans never use. However, this comparison, as far as relative clauses are concerned, ignores the fact (mentioned above) that it is not a two-way choice, but at least a four-way choice (five in Shakespeare) for relatives which are objects of the verb, six or eight for prepositional objects. Let us compare all of these, distinguishing again between restrictive and non-restrictive clauses. The four restrictive forms mentioned for verbal objects in Householder 1969 were illustrated by

- (1) The woman who Peter hated. (MIT preferred form)
- (2) The woman whom Peter hated.
- (3) The woman that Peter hated.
- (4) The woman Peter hated.

For Shakespeare and the Bible (as well as Chaucer) we must add

- (5) The woman which Peter hated.

Replace "the woman" by "Mrs. Jones," for non-restrictive types. For prepositional objects I listed six subtypes:

- P_1 The woman who Peter talked to.
- P_2 The woman whom Peter talked to.
- P_{1A} *The woman to who Peter talked.
- P_{1B} The woman to whom Peter talked
- P_3 The woman that Peter talked to.
- P_4 The woman Peter talked to.

There is no nonrestrictive for P_4 . There is a further subtype of prepositional

relative, the partitive construction (virtually always nonrestrictive) which does not offer as many choices :

PP_{2c} The women, five of whom Peter hated

PP_{1B} The women, of whom Peter hated five.

[PP₁ "The women who Peter hated five of" and the other three with *whom*, *that* and zero replacing *who* seem to be out because they would be restrictive.]

The speakers who claim that they never in their lives use *whom* (see Jespersen 1940, 483, quoting Sweet ; I have had several students who made this claim) cannot be telling the truth unless they never use such a partitive sentence (and, in fact, these sentences with "some of whom" "none of whom", "a few of whom", and the like are not rare).

Here, then are tables first for the verbal objects, then for the prepositional objects, again adding *which* to the list for Shakespeare and *The King James Version*.

Table 2
Objects of verbs

	Restrictive					Nonrestrictive			
	who	whom	that	0	which	who	whom	that	which
Shakespeare	3	86	71	52	151	16	144	5	13
Bible	0	11	3	0	3	0	7	0	0
Old Notes US	9	8	n.d.	n.d.	n.d.	2	0	n.d.	n.d.
Old Notes Br.	18	14	n.d.	n.d.	n.d.	9	1	n.d.	n.d.
US Myst.	1	6	4	56	1	1	2	0	0
Br. Myst.	1	4	4	52	0	1	3	0	0
Brown U	0	21	5	n.d.	0	1	20	0	0
Quirk	4	9	10	14	1	0	2	1	0
Algren (dialog)	0	1	0	7	0	0	0	0	0
Higgins	1	12	86	77	0	1	4	2	1
Lington	0	0	2	83	0	0	7	0	0

(n.d.=no data available)

Several things are obvious here :

(1) for restrictive clauses, zero is heavily favored by everyone modern except Higgins, who has *that* a few more times. This may be a local Boston dialect trait, or something idiolectic with Higgins, who also has *that* far more often as human subject than anyone else since Shakespeare. (I don't have count for Mark Twain, but one text of his also shows a high incidence of *that* as human subject and object).

(2) Only my old detective novel counts (both British and American) show more instances of *who* (restrictive or not) than of *whom*, and British novels alone show a substantial excess of non-restrictive *who* over *whom*. I believe that this corresponds well to American educated usage (what you overhear on

campus and on television or radio), but there is no American count, and Quirk's British figures (0 to 2) are statistically useless. We need a corpus ten times the size of his. It clearly does not characterize the dialects represented by Algren, Higgins and Linington. Shakespeare shows a substantial dominance of *whom*, followed by *that* and zero. The Brown University corpus's excess of *whom* may reflect American editorial practice. *Which* in modern usage, outside of a few regional dialects, occurs with human antecedents only if they can be represented by *it*, e.g. *committee, group, team, union, board* and the like, and are not counted here, although there are a few occurrences (5) in Higgins of real *he* or *she* type antecedents.

The variation among the five object types in Shakespeare (four elsewhere) is not quite as free as one might think, in particular as concerns *who* and *whom*. It is possible to find stylistic alternation (either (1) in conjoined relatives "The one who did X and which did Y", (2) stacked relatives "the one who did X that did Y", (3) cleft sentences "It's the man who does X that does Y" or (4) parallel sentences "Great is the one who did X and great is the one that did Y") involving most pairs: (1) *That* and *whom* or *who* (*Cymbeline* 1.6.144-8, *King Lear* 2.2.17-24, *Coriolanus* 2.2.9-11, *Pericles* 3.2.92, etc.); (2) *who* and *which* (*Henry V* 1.2.65, *Macbeth* 5.1.55, *Tempest* 5.1.160); *that* and *which* (*Henry IV B* 1,1.98-9, *King Lear* 2.1.62-4); Higgins *Choice* p. 176 "There's some people that's bad off which the state has got to take care of"; but no alternations of *who* (object) and *whom*.³ It would either be ungrammatical or stylistically inconsistent to say "That's a man who we like and whom you admire."

The table for prepositional objects has few surprises:

	Restrictive						Nonrestrictive			
	P+who	who,,P	P+whom	whom,,P	that,,P	0,,P	Pwho	who,,P	Pwhom	whom,,P
Shakespeare	0	0	46	7	11	17	0	0	81	4
U. S. Mysteries	0	0	5	0	1	15	0	0	10	0
British Mysteries	0	0	6	1	1	16	0	0	14	2
Brown U.	0	0	38	1	4	n.d.	0	0	42	0
Quirk	0	0	5	0	4	4	0	0	4	0
Algren (dialog)	0	0	18	0	0	8(5)	0	0	1(0)	0
Higgins	0	2	13	0	22	40	0	0	0	0
Linington	0	0	4	0	1	42	0	0	0	0

We find a categorical rule for relatives, that if the preposition immediately precedes (P *whom*) it is always *whom*, and when it dangles, we usually

³ Jespersen 1949_a, 87-92 lists quite a few examples, including *which-who*, *that-which*, *that-who*, *zero-that*, *zero-whom*, *zero-who*, and *who-that*.

have zero, though Shakespeare, Higgins and Quirk have a surprising number for *that*, which has merely trace representation elsewhere. The zeros for *who* go all the way. There are a small number of instances of *which* (both restrictive and non-restrictive with both initial and final prepositions) in Shakespeare, and even one in Higgins (*A year or so* P. 107 “in the small number of which [sc. drinkers] I am proud to count myself”). It is interesting that the first *wh*-relative for humans, *which*, is now almost extinct while the second, *whom*, is still strong and vigorous, being categorical after prepositions and frequent (especially if non-restrictive) elsewhere. *That*, which is still frequent in nonrestrictive use (subject, object, prepositional) in Shakespeare, almost disappears in recent texts.

The Bever and Langendoen article mentioned earlier also sets out historical stages, but these do not depend on case-marking, as Klima's do, and on most points the stages seem to be justified, though they fail to consider the fact that *whom* appears commonly as a relative centuries before *who* does. But two matters which they do include are worth our consideration: (1) the occurrence of zero subject relatives, and (2) the occurrence of 'returning pronouns' inside relative clauses, e.g. "his" in "This is the man that we saw yesterday and *his* wife was wearing red". B. and L. give stage 6 (1900 on) as marking zero subjects ungrammatical, and stage 4 (1550-1700) as making returning pronouns ungrammatical. Our modern corpora are all stage 6, and Shakespeare is stage 4. Let us first look at the occurrence of zero subject relatives. The choices are fewer than with objects.

Table 4
Human Subjects

	Restrictive			Nonrestrictive			
	who	that	0	which	who	that	which
Shakespeare	120	<u>1622</u>	79	49	<u>331</u>	268	49
Bible (A.V.)	0	89	0	21	8	3	19
US Mysteries	<u>330</u>	25	2	0	<u>117</u>	0	0
British Mysteries	<u>348</u>	11	1	0	<u>115</u>	0	0
Brown U.	<u>1473</u>	42	n.d.	0	<u>590</u>	1	0
Quirk	<u>200</u>	19	1	0	25	0	0
Algren (dialog)	<u>242</u>	24	24	0	28	0	0
Higgins	369	<u>372</u>	13	0	4	3	1
Linington	<u>329</u>	12	25	0	<u>116</u>	1	0

One clear stage boundary, after Shakespeare and the Bible, is the change in relative frequency of *who* and *that*--4 to 1 *that* in Shakespeare, 11 to 1 in the Bible, and then between 3 to 1 and 40 to 1 *who* in everyone else except Higgins, who has almost equal numbers of *who* and *that*. Clearly there is no categorical rule change, but there is a stylistic change. Similarly *which* is favored 5 to 1 over *who* in the Bible, and appears in respectable numbers in Shakespeare, but almost disappears later. The example in Higgins sounds authentic to me.

(*A year or two* p. 117 "The flower girls and Jews for Jesus, which wouldn't be parading around...") All zero relatives are restrictive from the earliest times. Among the moderns *that* becomes mainly restrictive; there is one nonrestrictive *that* in Linington, and 3 in Higgins, all of which may be considered colloquial or substandard in some way.

As for returning pronouns, they are too few to put into a table, but they occur from Shakespeare to Higgins. Here are a few from Shakespeare: (with *his*, which is also the most frequent today, I believe) *Hamlet* 1.2. 187-8 "He was a man I shall not look upon *his* like" (a zero which is neither subject nor object) and *Tempest* 3.3.30-2 "There are people of the island, who, though they are of monstrous shape, yet, note, *their* manners are more gentle-kind" and, a little further on, 3.3. 53-6 "you are three men of sin, whom Destiny...the never-surfeited sea hath caused to belch up *you*." In Higgins, note *Choice of Enemies* 57: "Mrs. Benedict, that was just on the stand this afternoon and Feldman got *her* all balled up." (Another non-restrictive *that*.) These may well be colloquial, but I would hesitate to call them ungrammatical.

Finally, this table shows the occurrence of zero subjects, which Bever and Langendoen outlaw for stage 6, even for existential and cleft sentences, and for all stages in other sentences. It is true that even in ME times zero subject relatives occur often in existential sentences (e.g. modernized from Steinki 85, "There is no man will give so much"), and certainly that is true in Shakespeare and true today, though less so than in Shakespeare.

Table 5

	There's	-have-	cleft	other* "be"	other
Shakespeare	25	5	18	16	15
Algren	2	0	10	8	4
Higgins	3	1	2	3	4
Linington	2	0	6	13	4

Note*: These are mainly of the type "He's the one..." or "She's the kind."

It is clear that existentials predominate only in Shakespeare; for all four combined, the "other *be*" category is the largest. Bever and Langendoen in their note 11 cite two passages from Shakespeare (*Cymbeline* 1.1.105 and *Pericles* 1.1.135) and quote Curme with approval when he suggests S. was deliberately being ungrammatical (it's not clear why). But among the 13 other examples in Shakespeare are several very good ones (which the *Cymbeline* passage is not, if my edition is followed in putting a semicolon after 107). And all three of our other authors have good examples, and Cheshire also cites two from her Reading teenagers. For instance: *Titus Andronicus* 1.1.267-8 "He () comforts you can make you greater than the Queen of Goths." Richard II 2.2. 127-8 "Our nearness... is near the hate of those () love not the king." *Hamlet* 4.7.130 "we'll put on those () shall praise your excellence." *Timon of*

Athens 1.1.100-1 "Your honorable letter he desires to those () have shut him up" 3.4.105 "Many () do keep their chambers are not sick" *Pericles* 1.4.74 "Thou speak'st like him ()'s untutored to repeat" And even one with a zero relative conjoined to a participle: *As You Like It* 2.4.70-1 "Here's a young maid with travel much oppress'd and () faints for succour."

Algren, *Man with the Golden Arm* 16 "You catch the guy () sliced up the little girl", 143 "I married the biggest dummy () ever walked in shoe-leather", Linington (as Shannon), *Knave of Hearts* 256 "I found the feller () sold it to him", 139 "We had him spotted for...somebody () comes to the beach just weekends", Higgins, *Choice of Enemies* "I kept fighting off the publisher () had this very talented nephew." Cheshire 74 "Owen Kelsey () used to live next door to us has emigrated." "Tell us the one about the lady () couldn't get a lift."

I think there is little doubt that in modern times the construction is colloquial, but it's certainly alive; in Shakespeare it even seems to be a little grand--clowns and rustics aren't the ones who use it. It is interesting that three of the examples quoted above have *those* as antecedent, possibly a faint trace of earlier days when this pronoun could function as a relative.

Before we leave our statistics entirely, it is worthwhile noticing the relative frequency of subject relatives and object relatives of all types, and of restrictive relatives compared with nonrestrictive ones. In this table I will lump together non-subject uses--object, object of preceding preposition, object of stranded preposition, and omit all uses of *which* and the nonrestrictive uses of *that*, which are significant only in Shakespeare.

Table 6

	who(m)				that		zero	
	S	O	NS	NO	S	O	S	O
Shakespeare	120	146	331	242	1,622	82	79	69
Bible	0	12	8	8	89	3	0	0
US Mysteries	330	11	117	13	25	5	2	71
Br. Mysteries	348	12	115	19	11	5	1	68
Brown U.	1,473	16	590	62	42	7	n.d.	
Quirk	200	18	25	6	19	16	1	18
Algren	243	4	28	1	24	0	24	44
Higgins	369	28	4	6	372	108	13	117
Linington	329	3	116	7	12	2	24	125
Totals	3,412	295	1,334	364	2,216	228	144	512

All Subjects=7106

All Objects=1399

This is a lopsided preponderance of subject relatives.

It is common to consider cases of variation as linguistic change in progress; but of course all languages are in process of change at most points continuously. However, change is not necessarily very fast. What changes can we see

here?

One change took place well before Shakespeare, the replacement of the three OE relatives (including a declinable pronoun whose neuter singular was *that*) by a single indeclinable *that*. A second which was substantially over shortly after Shakespeare (remember, we have no data⁴ between Shakespeare and 1900 or so--mostly 1950 or so) is the introduction of (human) *which* (before Chaucer's time) and its virtual disappearance between Shakespeare and today (probably before 1800, to judge by Jespersen 1949, 119). We also lack the beginnings of the introduction of *who* and *whom*, which we have already discussed briefly, nothing that *whom* came first, especially with prepositions. The same passage of Jespersen also mentions the interesting point that in the King James Bible (or Authorized Version, AV) human *which* is not used with prepositions, but only *whom*. Although AV and Shakespeare are nearly contemporary, AV is definitely archaizing and Shakespeare relatively colloquial. Still, even in Shakespeare, restrictive *who* is far outnumbered by *that* (table 4) though restrictive *whom* slightly outstrips *that* (Table 2). For us today, as tables 4 and 2 show, restrictive *who* and *whom* substantially outnumber *that* except in Higgins. Romaine 216 indicates that some modern Scots never use *who* or *whom* (or *which*), but only zero and *that*. It is lucky that we have Higgins in our corpus, or we would not realize that the race between these two is still going on. This race (which in fact involves the replacement of *that* by *who/whom*--see Jespersen 1949a, 80 "a growing tendency to extend the sphere of the *wh*-words") was already viewed the other way in the 18th century by Steele (as cited by Steinki 24; Jespersen says Addison--anyway Spectator no. 78), as some modern writers⁵, too, view the matter, whether as historical development or transformational derivation. To sum up, the sequence of relatives referring to human antecedents (substitutable in the singular by *he* or *she*) is approximately as follows:

After 1050 *that* and *se* (declinable);

1150 *that* and perhaps zero subject;

1300 *that*, *which*, *the which*, *which that*, *whom*, zero subject;

1500 zero subject and object, *that*, *which*, *the which*, *whom*, *who*
(at first only nonrestrictive)

All these last occur in Shakespeare though *the which* is a little scarce (6 or 7 examples).

The first question addressed by recent writers on this subject is that of conditioning factors: what contextual (syntactic or phonological) facts encourage or discourage the choice of one of the available alternatives? What sociological or psychological factors do likewise? Quirk and Romaine are

⁴ Jespersen 1949a, 84 gives some figures for eight 19th century writers, but the sample is small, and he does not separate objects from subjects and none of the sample is colloquial.

⁵ E.g. Sapir 174, a "drift" toward *that* and zero instead of *who* or *whom*.

especially keen on these points, though people like Steinki may speak of popular or poetical or literary prose usage, and even of conversational style, and Jespersen, in particular, gives attention also to factors like "after interrogative pronouns", "[after] superlatives and similar words", "those introduced by *it is*" [i.e. cleft sentences], not to forget "the desire to avoid monotony." Of all the proposed conditions, one that we have been using throughout this paper is the most clearly grammatical, the distinction between restrictive and nonrestrictive clauses. This condition (restrictiveness) has been categorical for zero relatives since the beginning, e.g. Chaucer's *Prologue* 529 "with him there was a plowman was his brother", though Mustanoja (121,205) claims that this and many other sentences like it are really deleted personal pronoun subjects. And, as table 2 shows, it has become nearly categorical for *that*, though Quirk has one and Higgins two nonrestrictive *that*'s (and Romaine's modern Scots must have many). The only other categorical rule visible in the data is position immediately after a preposition; this guarantees *whom* in modern relative use, though *which* is possible earlier, and almost guarantees *whom* for questions (unless nothing else follows--"with who?" reclamatory, as a rule).

Quirk lists the following additional conditions:

1. Human (he or she) or not.
2. follows the antecedent immediately or not.
3. clause is or is not initial or final in the sentence.
4. antecedent is complex.
5. subject of relative clause (with relative object) is or is not a personal pronoun.
6. number of words in clause (not counting the introductory word).

Cheshire repeats Quirk's 1 and 6 and adds one:

7. is there a demonstrative *that* in the preceding clause?

Romaine considers also whether the antecedents are

8. definite or not,
9. qualified or not,
10. superlative or not (Jespersen also mentions this).

She also specifies four stylistic categories for her Scottish data: official prose, non-official prose, epistolary prose and verse.

I have made many attempts to detect categorical rules based on some of these criteria, particularly for Shakespeare. Here are some results.

The criterion of the position of the antecedent in its own clause, whether or not it precedes the main verb or not doesn't help much: Zero object ("The man I saw" clauses have pre-main verb antecedents 23% of the time in Shakespeare, *that* clauses ("The man that I saw") 25%, and *who* or *whom* clauses ("The man who(m) I saw") about 34% of the time. Hardly categorical. Of course by the simpler criterion "antecedent immediately precedes", zero clauses (both S and O) are nearly 100% in Shakespeare, as today also.

The criterion of number of words in Shakespeare is marginally interesting.

Table 7

	zero-(0)	that	who(m)
4 or less	81%	58%	54%
5 or more	19%	42%	46%

	4 or less	5 or more
Zero	33%	13%
That	32%	39%
Who(m)	35%	48%

Still, this only says that zero-object clauses are likely to be short, not that short clauses are likely to be zero.

I also counted for the criterion "Is the first word of the clause an unstressed subject pronoun (*I, he, she, etc.*)?" with similar results.

Table 8

	zero	that	who(m)
+Pronoun	93%	70%	53%
-Pronoun	7%	30%	47%

	+Pronoun	-Pronoun
Zero	36%	6%
That	38%	38%
Who(m)	26%	56%

This is a little better. Perhaps if I had multiplied the two criteria, i.e. counted clauses which are both under 4 words and initiated by a pronoun, I might have got something nearly categorical.

The other one I checked for was number (singular or plural) of antecedent.

Table 9

	zero	that	who(m)
Plural	31%	37%	29%
Singular	69%	63%	71%

	Plural	Singular
Zero	27%	28%
That	45%	36%
Who(m)	29%	35%

This looks even less interesting.

We have already mentioned some other good conditioners for *who*: virtually all continuative (connecting) relatives are introduced by *who(m)* or *which*, and so are all indefinite-general type clauses in which the whole relative clause precedes the main clause (with or without an "antecedent" or resumptive pronoun). The [+human] test applies nowadays perfectly to *who*

and *which*, though in Shakespeare *which* may be human and *who* inanimate (not always easily interpretable as personification, e.g. in *The Merchant of Venice* 2. 7. 4 where the first of the three caskets (of gold) is "who this inscription bears", the second (of silver) "which this promise carries" or *Troilus and Cressida* 3. 3. 201 "There is a mystery with whom relation durst never meddle..., which hath an operation.,,")

I checked all of the proposed conditioning factors except "complexity of the antecedent" and could find no substantial biases for or against any one of the three choices, especially in view of the small numbers involved (Cheshire's are even smaller, of course), except possibly that a quantified antecedent (*all, each, every, no, none, any*, etc.) slightly favors *that* over *who* and zero. Even this is certainly not categorical.

What about stylistic factors, what I have elsewhere (Householder 1972, 1983) called labels? Since it is here that the essence of linguistic change lies, in my view, I should naturally try my best to identify the ones involved. The difficulty is finding objective evidence for these labels; even in contemporary usage people may disagree. Apparently MIT linguists regard "The man who I saw" as normal U.S. educated colloquial, whereas to me it seems like pretentious pseudo-colloquial. But in contemporary use a few things are more or less agreed on: much of Klima's L_4 , if combined with other similar usages, would be judged low-class uneducated style ("Him and me ain't no stupid jerks, nohow" for instance), though other L_4 examples seem as good as L_3 ("It is me"), e.g. "Is she as tall as me?" (from Shakespeare) or in answer to "You did it!" "who, me?"

It is also not hard to find examples which most people would agree are extremely formal or hypercorrect, e.g. most instances of interrogative *whom* without a preposition before it ("Whom did madam wish to see?") And there are instances of both sorts in the language of our detective novels and other modern material. Higgins wants to depict one of his characters as hypercorrect, and makes him say things like this: "I asked Sam ...why an experienced politician such as I should accept... commands from two novices such as they." (*A City on a Hill* p. 204) A lot of the characterization in this case depends on the use of learned words and long periodic sentences, but bits of L_1 now and then help out. But a lot, perhaps most of this 20th century material lies somewhere in between. And when we turn to Shakespeare, the task is even more difficult. In cases like what I have called $L_{2,1/2}$ the unanimity of the testimony shows that Shakespeare did not regard this style ("between you and I" etc.) as either high or low or affected, but as a norm, and the exceptional cases (e.g. in *Antony and Cleopatra*) must then be a bit more colloquial. But most of his pedants and malaprops are characterized by long words, involved prose and what they talk about, rather than by their choice of pronouns or cases. I think there is no question that $L_{2,1/2}$ now arouses reactions like Kilpatrick's (quoted above), while much L_4 passes as rather normal; so here

we can surely identify a linguistic change. But in most cases we cannot.

A final question may be asked about the nature of the variation and of the incipient or potential change involved: is it syntactic? or is it lexical? Both Klima and Bever-Langendoen seem to regard it as syntactic, and assume that there is a rule of "case attachment" or the like involved, and propose various other rules and rule-orderings. But the only reason I can see for not treating it as lexical is that one of the lexical variants is a phonological zero. Can we have a lexical entry with no phonological substance? In years gone by much was written about zero morphs or morphemes, but certainly things like *t* and *e* or even COMP, in some cases, provide a reasonable justification for entering a relative particle zero, specified as object-deleting (or the like) and neutral as to gender and number, and perhaps another one which is subject-like and has some other restrictions.

Most of the current treatments also assume a single underlying relative (*who* for [+human]) which is then replaced by *whom* or *that* under certain conditions, and is sometimes deleted, either early or else after being changed to *that*. If these are all partial synonyms, do we say that underlying both surface *thesis* and surface *dissertation* is one or the other of the two, which is then changed by rule? Various attempts to be more scientific and precise about variation are exhaustively dealt with by Romaine, and I feel she has found them all unsatisfactory, while implying that it may be possible to come up with a satisfactory theory.

I believe that labels are real, even though only partially shared by members of a speech community. And certainly labels change. But in the case of the human object relative clause it is truly amazing how little has changed since Shakespeare.

REFERENCES

- Bartlett, J. (1906) *Complete Concordance to... Shakespeare*, MacMillan, London.
- Bever, T. G. and D. T. Langendoen (1969) 'The Interaction of Speech Perception and Grammatical Structure in the Evolution of Language,' in Stockwell, Robert P. and Ronald K.S. Macaulay, eds., *Linguistic Change and Generative Theory*, Indiana U. Press, Bloomington.
- Cheshire, Jenny (1982) *Variation in an English Dialect*, Cambridge U. Press, Cambridge and New York.
- Householder, Fred W. (1969) Review of *Language and its Structure* by R.W. Langacker, *Language* 45, 886-897.
- (1971) *Linguistic Speculations*, Cambridge U. Press, Cambridge and New York.

- (1972) 'The Principal Step in Linguistic Change,' *Language Sciences* 20, 1-5.
- (1983) 'Kyriolexia and Language Change,' *Language* 59, 1-19.
- Jespersen, Otto (1949a) *A Modern English Grammar*, Part III, Syntax, Second Volume, George Allen and Unwin Ltd., London.
- (1940) Part V, Munksgaard, Copenhagen.
- (1949b) Syntax Part VII, Munksgaard, Copenhagen.
- Klima, Edward S. (1964) 'Relatedness between Grammatical Systems,' *Language* 40, 1-20.
- Kučera, Henry and W. Nelson Francis (1982) *Frequency Analysis of English Usage*, Houghton Mifflin, Boston.
- Lightfoot, David W. (1979) *Principles of Diachronic Syntax*, Cambridge U. Press, Cambridge and New York.
- Quirk, Randolph (1968) *Essays on the English Language, Medieval and Modern*, Indiana U. Press, Bloomington.
- Romaine, Suzanne (1982) *Socio-historical Linguistics*, Cambridge U. Press, Cambridge and New York.
- Sapir, Edward (1939) *Language*, Harcourt, Brace and Company, New York.
- Schmidt, Alexander (1875) *Shakespeare-Lexicon*, Reimer, Berlin.
- Spevack, Marvin (1969-70) *A Complete and Systematic Concordance to the Works of Shakespeare*, Vols. 4, 5, 6 *A Concordance to the Complete Works*, Georg Olms, Hildesheim.
- Steinki, Johannes (1932) *Die Entwicklung der englischen Relativpronomina in spätmittelenglischer und frühneuenglischer Zeit*, Breslau 1932 (Inaugural-Dissertation).
- Stockwell, Robert P., and Ronald K.S. Macaulay (1972) *Linguistic Change and Generative Theory*, Indiana U. Press, Bloomington.
- Tatlock, J.S.P. and A.G. Kennedy (1963) *A Concordance to the Complete Works of Geoffrey Chaucer*, Peter Smith, Gloucester, MA.

Department of Linguistics
 Indiana University
 Bloomington, Indiana 47405
 U.S.A.