

## 1 Introduction

### 1.1 Discourse markers

Many languages use *discourse particles* to convey a Speaker commitment or attitude towards the proposition expressed, without contributing to descriptive content (Zimmermann, 2011). While English is thought not to have discourse particles proper, it does possess parallel items which convey similarly rich Speaker commitments, for example for *man* in English (McCready, 2008). I'll argue that intersentential *what* is another example.<sup>1</sup>

#### Plot.

DESCRIPTION — Illustrate phenomenon with examples.

CORE PROPERTIES — Major syntactic and prosodic patterns.

FORMAL ANALYSIS — Provide general arguments against two intuitive accounts: (a) the rising declarative and (b) the syntactic *wh*-in-situ question account, and propose an approximative account.

#### Plot spoiler.

I'll propose that intersentential *what* raises issues, like a question, through an approximative meaning (a la Lasersohn's 1999 pragmatic halos) which either can be resolved through a direct answer or left as an open issue depending on the Question under Discussion (Roberts, 1996). The compositional meaning captured as a set of propositions within an Alternative Semantics framework (Rooth, 1985; Kratzer, 1991).

<sup>1</sup>Many thanks to the participants at CUSP 5 in UC San Diego, Jay Atlas, Noah Constant, Donka Farkas, Lyn Frazier, Angelika Kratzer, Chris Potts, Jessica Rett, Lisa Selkirk, Radek Šimik, and SALT reviewers for their feedback!

### 1.2 Intersentential what

Typical examples from a variety of genres:

- (1) There's been what? A dozen and a half murders since I've been here. (Hammett, 1929: *Red Harvest*)
- (2) I recognize you though it's been, what, 30 years? (Midsomer Murders *Death's Shadow*, 1999)
- (3) You think that what, twenty deserters from the Sudanese army are going to come back and make Sudan a Communist nation? (Eggers, 2006: *What is the what*)
- (4) Woody's what? 73? he'll be writing this stuff 'til he's what? 95? so we might as well get comfortable ...

Comments on Popwatch about a Woody Allen film, 2009

<http://popwatch.ew.com/popwatch/2009/05/whatever-works.html>

### 1.3 Terminology

Ignore sarcastic and aggressive uses of *what* which often appear preceded by *so* in sentence initial position (Dehé and Kavalova, 2006).

- (5) So, what, my father was Dr. Mengele? (*Heroes*: Season 3)

Reserve what to demarcate cases of interest and use *what* for other string identical cases.

#### A term of art.

COMPLEMENT — Constituent following what

- (6) She'd only been governor for, what, two years .  
complement

## 2 Core properties

### 2.1 Syntactic licensing

Complement of what syntactically unconstrained:

Number phrases:

- (7) a. You've been here for, what, two years? (DP complement)  
 b. You've been here, what, for two years? (PP complement)  
 c. You've what, run 12 marathons so far? (VP complement)

Also appears with non-number complements, contra Dehé and Kavalova (2006), in which what is licensed by syntactic feature matching with an abstract NUM feature.

Non-number phrases:

- (8) It's filled with what, whip cream ... and strawberries and something (Switchboard)  
 (9) I haven't seen you since, what, Lebanon? Bosnia? (Burn Notice: Season 3)

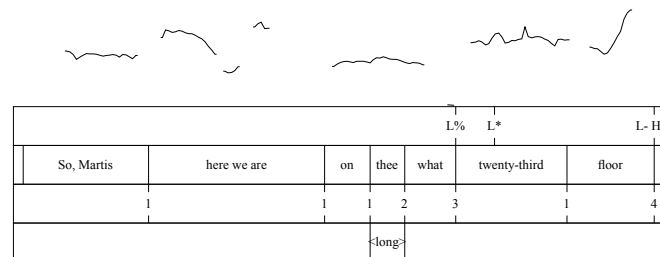
Non-number complements are not marginal: in a mini-analysis of Switchboard of what (18 instances of approximative what from first 2,000 cases of what), 40% involved non-number complements (Appendix A).

### 2.2 Prosodic realization

- Produced with a low (L%) or low-fall (H- L%) boundary tone
- Often offset from its host by pauses on the right boundary
- Complement is typically – but not always – accompanied by a rising, question like tune at the end
- Assume that what syntactically F-marks an element in the complement; almost always immediately adjacent element.

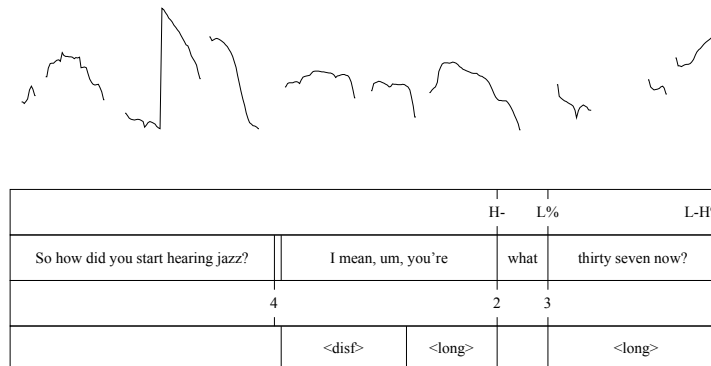
- (10) So, Martis, here we are on the, what, twenty-third floor?

(BBC)

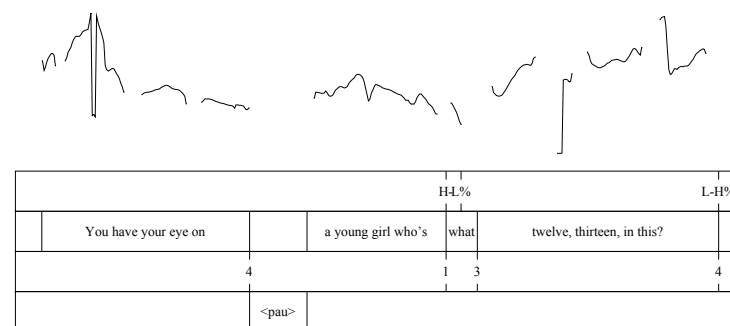


(Hat tip to Radek Šimik for this example)

- (11) I mean, um, you're, what, thirty-seven now? (NPR)



- (12) Though I will say, you know, you have your eye on a young girl who's, what, 12, 13 in this? (NPR)



### 3 Flavors of analysis

#### 3.1 Rising declarative [rejected]

✗ Distributional mismatch: what conveys Speaker’s uncertainty, not Addressee’s public commitment.

#### 3.2 *Wh-in-situ* question with best guess answer [rejected]

✗ Doesn’t pattern with other *wh-in-situ* cases.

✗ Different prosodic and discourse constraints than *wh-in-situ* questions.

✗ Lack of syntactic correspondence with genuine question

#### 3.3 As approximator

✓ Gives rise to question meaning in specific contexts, otherwise approximative.

✓ Compositional derivation of propositional alternatives, through pragmatic halos.

#### 3.1 As rising declarative

(13) **Contextual bias condition.** Rising declaratives can only be used as questions in contexts where the Addressee is already publicly committed to the proposition expressed. (Gunlogson, 2001)

(14) A: There’s a leopard in the living room.

B’s response:

a. ? Is there a leopard in the living room? (Polar Q)

b. There’s a leopard in the living room? (Rising D)

c. # There’s what, a leopard in the living room? (what-marker)

(15) A: I know what this is (smugly holding up fruit).

B’s response:

a. Is that a persimmon? (Polar Q)

b. # That’s a persimmon? (Rising D)

c. That’s what, a persimmon? (what-marker)

The what-marker seems to pattern with (polar) questions, perhaps it is a question after all . . .

### 3.2 As *wh-in-situ* question

Intuitively, we might treat what as composed of two fused clause types: a *wh-in-situ* question, followed by best guess answer.

(16) You’ve been here what?      Two years?  
wh-in-situ question      best guess

**Issue 1:** Usually, *wh-in-situ* licensed in English when there’s another *wh*-element (e.g., Cheng, 2009). However, what is markedly odd with another *wh*-element (17). Such an analysis would depart from the normal syntactic analysis of English questions.

(17) a. Who said John drank what?

b. # Who said John drank, what, two beers?

**Issue 2:** Best guess answer would be mandatory, again departing from normal conditions on questions. Situations without answer are very limited, and have different prosodic realizations on *what*.

*I. Echo questions:* Addresses previous utterance, Speaker uncertain about what was said.<sup>2</sup>

(18) Waitstaff repeating order in a noisy bar:

You want a what? A beer? (Rising tune on *what*)

*II. Leading questions contexts:* Addressee expected to know the answer, no Speaker uncertainty.

(19) Teacher, somewhat condescendingly, to a student:

Now, students, four plus four is what?

<sup>2</sup>In a pilot study with Chris Potts, I elicited readings from two actresses performing sentences like *You want what two beers?* in (i) echo and (ii) lack of knowledge contexts. In echo contexts, what was given a rising prosody with increased intensity. In lack of knowledge contexts, what received a low or high-low fall. Subjects on Amazon Mechanical Turk confirmed that these prosodies were appropriate for the context.

**Issue 3:** Expect that *wh*-in-situ and canonical questions should be syntactic variants, yet they are not:

- (20) a. You're going to move in, what, 4 months?  
 b. # What are you going to move in? (4 months?)  
 c. When are you going to move? (4 months?)

Nevertheless, what likely introduces a question-like *issue* into the context, modeled as a set of propositions, cf. Hamblin meaning for questions.

### 3.3 As approximator

There are many cases in which the issue raised by what does not address the main point of the discourse. In these cases, the complement of what provides an approximately correct value, which goes unaddressed (unless the estimate is egregiously wrong).

- (21) [Message left on answering machine.]  
 Hello, it's Caroline Todd again. It's now 7:09 AM, so that's what, 8 and a half hours since you came to on your home on your way home from the Pink Lagoon.  
 (Green Room: Season 1)

Such uses put what on par with interjective fillers (Clark, 2004; Fox Tree, 2010), many of which linguistically signal the epistemic state of the Speaker (Clark and Fox Tree, 2002).

#### Conventional meaning of what.

The Speaker is not committed to the exact value of the complement, but instead *approximates* the value from a set of relevant propositional alternatives.

The strategy is to develop a compositional approach which honors this conventional meaning, but which *invites* responses to the issue raised, if it is pertinent to the Question under Discussion.

### 3.4 Formal account

**Step 1** Use tools from Alternative Semantics to produce a set of propositions constituting a *partial answer* to an issue raised by utterance with what.

**Step 2** Relation to QuD determines whether the Addressee is expected to further resolve that issue (question use) or leave it unresolved (approximative use).

#### 3.4.1 Step 1: Deriving relevant propositions

##### Sentence operator

The element what is a sentence operator  $\text{AROUND}_C$  which takes the ps-skeleton of a proposition and returns a set of propositions obtained by evaluating the ps-skeleton at assignment functions from a contextually restricted domain  $C$ .

- (22) *Ps-skeleton*: F-marked constituents translated as designated variables. (Rooth, 1985)
- (23) *Two variable assignments*: (Kratzer, 1991)
- a. Interpret an ordinary variable  $v$  of type  $\tau$  :  
 $\llbracket v_\tau \rrbracket^{g,h} = g(v_\tau)$
- b. Interpret a designated variable  $\mathbf{V}$  of type  $\tau$  :  
 $\llbracket \mathbf{V}_\tau \rrbracket^{g,h} : h(\mathbf{V}_\tau)$
- (24) You've been here, what, two years?
- a. F-marked: you've been here [two]<sub>F</sub> years
- b. Ps-skeleton: you've been here  $\mathbf{V}$  years
- c. P-set:  $\{p : \exists h. [p = \llbracket (24b) \rrbracket^{g,h}]\}$
- $$= \left\{ \begin{array}{l} \text{You've been here 0 years} \\ \text{You've been here 1 year} \\ \text{You've been here 2 years} \\ \text{You've been here 3 years} \\ \text{You've been here 4 years} \\ \text{You've been here 5 years} \\ \vdots \end{array} \right\}$$

- The P-set gives us a huge space of possibilities, much like a question as the set of possible answers (Hamblin, 1973).
- We can treat the contribution of what as restricting that set via an approximation of answers that cluster around the value provided by the complement.

$$\left\{ \begin{array}{l} \text{You've been here 0 years} \\ \text{You've been here 1 year} \\ \text{You've been here 2 years} \\ \text{You've been here 3 years} \\ \text{You've been here 4 years} \\ \text{You've been here 5 years} \\ \vdots \end{array} \right\}$$

Restrict assignment functions  $h$  to those that deliver values for  $\mathbf{V}$  within a contextually restricted domain  $C$ .

**Slack interval** Let  $\mathbb{I}$  be the interval consisting of values (from a contextually determined scale  $S$ ) centering around  $c_n$ , the value given explicitly in the utterance by some index  $\pm i$ :

$$\mathbb{I} = \{x : c_{n-i}^S \leq x \leq c_{n+i}^S\}.$$

**Assignment cluster** Let  $C$  be a set of assignment functions  $h$  for designated variables  $\mathbf{V}$ , such that for any  $h \in C : h(\mathbf{V}) \in \mathbb{I}$ .

Take what to operate on the ps-skeleton of its complement  $p'$  returning a set of propositions obtained by substituting values for  $\mathbf{V}$  from all assignment functions  $h \in C$  :

(25) For proposition  $p$ , let  $p'$  be the ps-skeleton:  
 $\text{AROUND}_C(p') = \{p : \exists h. [p = \llbracket p' \rrbracket^{g,h \in C}]\}$

(26) You've been here, what, two years?

a. F-marked: you've been here [two]<sub>F</sub> years

b. Ps-skeleton: you've been here  $\mathbf{V}$  years

c.  $\text{AROUND}_C(\llbracket (26b) \rrbracket^{g,h \in C})$

$$= \left\{ \begin{array}{l} \text{You've been here 1 years} \\ \text{You've been here 2 years} \\ \text{You've been here 3 years} \end{array} \right\}$$

### 3.4.2 Step 2: Discourse function

As discussed, the resulting set of propositions is compatible with *multiple discourse uses*. Whether the issue raised by what is resolved depends on how important that issue is to advancing the discourse – whether the issue is related to the QuD (Roberts, 1996) or is on the Table (Farkas and Bruce, 2010).

**A prediction:** If the Addressee is not in the appropriate epistemic state to resolve the question, uses of what should be taken as approximations, and go unaddressed. Examples from Switchboard support this prediction.

*Addressee familiar about issue raised = Question*

- (27) 1 utt1: okay,  
 1 utt2: I missed a part of it  
 1 utt3: we're to talk about what, lawns and gardens  
 2 utt1: lawn and garden work and what you enjoy and what kind of work you do

*Addressee unfamiliar about issue raised = Approximation*

- (28) 164 utt3: I moved down here from Chicago, what, twenty-three years ago  
 165 utt1: {f oh, } uh-huh  
 166 utt1: – [ and, + ] to Lewisville –

## 4 Conclusion

Presented an account of interjective what with the components:

### Cheat sheet.

1. Conventional implicature signaling Speaker uncertainty
2. Raises issue which the complement partially resolves
3. Issue + partial answer modeled as a contextually restricted set of propositions derived from the ps-skeleton in Alternative Semantics.
4. Meaning is flexible enough to be treated as a question or an approximation, depending on discourse status.

But see Appendix B for more issues ...

## References

- Cheng, Lisa Lai-Shen. 2009. *Wh-in-situ*, from the 1980s to now. *Language and Linguistics Compass* 3:767–791.
- Clark, Herbert H. 2004. Pragmatics of language performance. In *Handbook of Pragmatics*, ed. Laurence R. Horn and Gregory Ward, chapter 16, 365–382. Oxford: Blackwell.
- Clark, Herbert H., and Jean E. Fox Tree. 2002. Using uh and um in spontaneous speaking. *Cognition* 84:73–111.
- Dehé, Nicole, and Yordanka Kavalova. 2006. The syntax, pragmatics, and prosody of parenthetical *what*. *English Language and Linguistics* 10:289–320.
- Eggers, Dave. 2006. *What is the what*. New York: Vintage Books.
- Farkas, Donka F., and Kim B. Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27:81–118.
- Fox Tree, Jean E. 2010. Discourse markers across speakers and settings. *Language and Linguistics Compass* 4:269–281.
- Groenendijk, Jeroen, and Floris Roelofsen. 2009. Inquisitive semantics and pragmatics. Presented at the Workshop on Language, Communication and Rational Agency, Stanford University.
- Gunlogson, Christine. 2001. True to form: Rising and falling declaratives as questions in English. Doctoral Dissertation, University of California, Santa Cruz, Santa Cruz.
- Hamblin, C. 1973. Questions in Montague English. *Foundations of Language* 10:41 – 53.
- Hammett, Dashiell. 1929. *Red harvest*. New York, NY: Vintage Crime/Black Lizard.
- Kratzer, Angelika. 1991. The representation of focus. In *Semantik: Ein internationales Handbuch der zeitgenoessischen Forschung*, ed. A. von Stechow and D. Wunderlich, 805–825. Berlin: de Gruyter.
- Lasersohn, Peter. 1999. Pragmatic halos. *Language* 75:522–551.

- McCready, Eric. 2008. What man does. *Linguistics and Philosophy* 31:671–724.
- Roberts, Craige. 1996. Information structure: Towards an integrated formal theory of pragmatics. In *OSUWPL volume 49: Papers in semantics*, ed. Jae Hak Yoon and Andreas Kathol. The Ohio State University.
- Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, University of Massachusetts, Amherst, Amherst, MA.
- Zimmermann, Malte. 2011. Discourse particles. In *Handbücher zur Sprach- und Kommunikationswissenschaft HSK 33.2*, ed. Paul Portner, Claudia Maienborn, and Klaus von Heusinger, 2011–2038. Berlin: Mouton de Gruyter.

## Appendices

### Appendix A: Switchboard results

Mini-corpus study conducted on Switchboard fragment distributed in NLTK, in which the first 2,000 from over 7,500 instances of *what* were classified. 18 clear examples of *what* were found, and classified according to (i) Syntactic category of complement, and (ii) Whether the utterance elicited a response from the other participant.

	<i>NP</i>	<i>PP</i>	<i>Number</i>	<i>Total</i>
<i>Response</i>	7	1	3	11 (61%)
<i>No response</i>	0	0	7	7 (39%)
	7 (39%)	1 (< 1%)	10 (60%)	18 (100%)

Table 1: Categorization of 18 instances of the *what*-marker found in the Switchboard corpus

## Appendix B: A few caveats and remaining issues

It is unlikely that what has an approximative *lexical meaning* of its own. Other *wh*-words can serve a similar function, although less productively:

- (29) a. John moved to Claremont, when, last month?  
 b. Mary corrected, who, Phil?

Many other ways to signal a similar discourse function:

- (30) a. You've been here, what, two years?  
 b. You've been here ⟨pause⟩ two years?  
 c. You've been here like two years?

The what-marker combines with other approximatives, disjunctions, and multiple complements:

- (31) A. How long has he been dead?  
 B. According to the coroner's report, about 24 hours.  
 C. That's, what, like 2 days after the break in at Skylar Wyatt's house? (*Numb3rs*: Season 2)
- (32) How can you hate *The Sting*? It like, what, took 20 Oscars?! (*Castle*: Season 2)
- (33) a. I'm no expert but, my guess is that a high rise at this location has gotta fetch at least, what, three, four hundred million? (*Numb3rs*: Season 5)  
 b. I mean, face it you can buy pharmaceutical grade cocaine for what, ten or twenty dollars an ounce (Switchboard)  
 c. We've got, what, two or three percent of the population – a tiny number of Americans – who are sincerely saying, 'Let us into this institution: this means everything to us.' (NPR on marriage equality)

Additionally, the complement can be a non-scalar element, for example:

- (34) [Upon seeing a book on the  $\lambda$ -calculus in a cafe:]  
 You're, what, a computer scientist?

These kinds of best guesses might be treated in terms of *ad hoc* scales according to who, in this case, which profession, is most likely to read a book on the  $\lambda$ -calculus.

## Appendix C

Ultimately, the treatment of what should link up with a more general framework for discourse management. I think that Inquisitive Semantics may be particularly well-suited to this task, as it predicts the existence of *hybrids* – combinations of informative and inquisitive (issue-raising) elements (e.g., Groenendijk and Roelofsen, 2009).

**Informative** A proposition  $\phi$  is informative just in case there is a possibility for  $\phi$  and a possibility that  $\phi$  excludes; proposes change to the common ground.

**Inquisitive** A proposition  $\phi$  is inquisitive just in case it has two possibilities; raise issues.

	Informative	Inquisitive
Question	–	+
Assertion	+	–
Hybrid	+	+
Insignificant	–	–

Thus, speech acts containing a what element are both *inquisitive*, in that they raise issues, and *informative*, in that it provides a possible partial answer which may sufficiently update the discourse, depending on the discourse topic.