Some viruses in the semantics[*]

Marcin Morzycki
Michigan State University

1 Introduction

This is a squib about weirdness — what it is, how to make sense of it, and whether we should resist making too much sense of it. It's about phenomena that are in a specific sense strange, complicated, formal, not especially child-friendly, associated with wavering judgments, and prone to inviting conscious self-reflection.[1]

Sobin (1994, 1997) proposed that certain syntactic processes should be regarded as what he called *grammatical viruses*, rules that operate outside the grammar. One prominent example is case in coordinate structures. The pronouns in (1) surface as accusative normally, but can surface as nominative in formal contexts:

(1) a. Clyde and $\left\{ \begin{array}{c} \text{I} \\ \text{me} \end{array} \right\}$ left.

   b. It was $\left\{ \begin{array}{c} \text{I} \\ \text{me} \end{array} \right\}$.

Following Emonds 1986, Sobin suggests that the nominative forms are not really part of the grammar of English. He proposes that they are instead the result of a pre-scriptive rule — consciously acquired, awkwardly half-internalized — that speakers strive to observe, with only intermittent success. The trouble speakers seem to have with the nominative forms contrasts starkly with how effortlessly we deploy the accusative. This is the first indication that these two case configurations don't have the same status. The true grammar of English generates only the accusative forms, and in that respect it's simpler than it might otherwise have needed to be.

If the syntax can be infected in this way, it stands to reason that the semantics might be as well. I'll argue that it may be. The potential viruses I'll examine are the

[1] I can't imagine why I thought it might be appropriate for a collection of squibs in honor of Kyle.

word *respectively*, factor/ratio phrases, and certain mathematical expressions such as *zero* (Bylinina & Nouwen to appear).

Sobin suggested thinking of viruses as, by definition, rules of the syntax, so the notion will have to be stretched slightly to extend it to the semantics. I'll also use the term "virus" itself perhaps more loosely than he intended, to include not just the principles that give rise to certain structures but also to particular lexical items and constructions themselves.

Of course, it's insufficient to simply announce that viruses don't need to be explained as part of the grammar proper. They still need *some* explanation. As Sobin says, without that, "'virus' becomes merely a label for unexplained phenomena". But even in the absence of a fully-developed theory of viruses, the benefit of recognizing viruses is, first, that doing so may account for empirical differences between viruses and the ordinary grammar, and, second, that it may simplify the grammar significantly.

2 Properties of viruses

Sobin (1997) suggest that viruses are distinguished by the following properties:

- They occur primarily in the prestige form of the language.

- They present a challenge for acquisition and are acquired relatively late.

- They may require what he calls "tutorial support" in the form of explicit prescriptive injunctions.

- Speakers have difficulty controlling them. They're prone to slip-ups and hypercorrection.

- They are consciously applied. Speakers make a self-aware attempt to observe a rule.

The case facts in (1) have all of these properties. *Clyde and I* is most natural in formal contexts. It's not the sort of thing young children say — but it is the sort of thing a schoolmarm might attempt to "correct" them with. The last point is perhaps the most revealing. Hypercorrect forms of the sort in (2) are widely attested:

(2)   a. She fed Alpo to Clyde and I.
        b. There have been some bitter disagreements between Clyde and I.

These are of course accusative positions, and the prescriptive injunction favors *me*. This involves over-application of the rule, but slip-ups in the other direction are

also natural. It wouldn't be hard to find speakers that strive mightily to observe it, in some cases as something akin to a moral obligation, and yet without conscious effort fail to apply it in contexts in which it's "supposed" to apply.

Sobin's principal other example involves agreement in existential constructions (but see Schütze 1999 for an argument against this). In (3), singular agreement is natural, but prescriptively disfavored:

(3)   There's fancy cufflinks on that suit.

Another probable virus is the nominative form in (4), which in most contexts would come off as almost comically pretentious:

(4)   It was I.

Lasnik & Sobin (2000) add the example of *whom*. Satisfyingly, the virus that gives rise to (4) reveals that *whom* is viral too. The only natural choice in (4) is of course the accusative *me*. *Whom* was historically an accusative form, but (5) is decidedly not natural:

(5)   *Whom was it?

Yet this is precisely what we would expect if it were straightforwardly accusative, for precisely the same reason that we normally use the accusative form in (4). Yet it seems safe to assume that no prescriptive injunction was ever required to stigmatize (5) and to compel speakers to use *who* in its place.

The analytical intuition Sobin and Lasnik articulate is that speakers engage in a kind of real-time self-monitoring, consciously editing their speech in accord with the demands of viruses.

3   "Respectively" and its kin

There is a not-so-small literature struggling with the lexical semantics of *respective* and *respectively* (Dalrymple & Kehler 1995, Gawron & Kehler 2004, Kac 1987, McCawley 1968, 1998, Stockwell et al. 1973, Moltmann 1992, Pullum & Gazdar 1982, Kay 1989, Eggert 2000, Okada 1999). The principal challenge involves sentences such as (6):

(6)   Fido and Felix ate Alpo and Whiskas respectively.

The principal challenge is that *respectively* flags that each element of the conjoined subject be paired with each element of the conjoined object in the order in which

they were mentioned. *Vice versa* behaves similarly (Kay 1989, Farkas 2015), as does the adjectival form *respective*.

This is odd and unnerving. It requires that the semantics make reference to the linear order of constituents in the sentence. We don't normally assume that it has this ability (though Morzycki 2008 recklessly raises this possibility in another context), but the conclusion seems virtually inescapable. Clearly, (6) doesn't have the same truth-conditions as (7):

(7)   Fido and Felix ate Whiskas and Alpo respectively.

Yet how to achieve this result? Superficially, this would seem to suggest that the denotation of *Alpo and Whiskas* should be different from *Whiskas and Alpo*. If they're directly referential, it's unclear how to accomplish this. It would seem to imply that there are two different plural individuals in the model that each of these (respectively) denote.

Alternatively, we could assume that these nevertheless do have the same denotation, and that *respectively* invites the addressee to, metaphorically, play back the tape of the utterance to match up the members of coordinate structures correctly. For what it's worth, my intuition is that this is precisely what happens, and in parsing these I feel a conscious awareness of the tape rewinding.[2] This, of course, requires that there be a tape in the first place, and that the semantics have access to it.

What Gawron & Kehler (2004) propose is roughly along these lines. They introduce a "sequencing function" as an element of the discourse context, which maps non-atomic objects of arbitrary type to assignment functions, which in turn map natural numbers to objects. A sequencing function might, for example, reach into the plurality of Felix, Fido, and Rover, and pull out the second-mentioned atomic member (Fido). The sequencing function, then, presupposes something like a tape of preceding discourse. Suppressing the necessary cross-categoriality and taking notational liberties, their denotation is as in (8), where f is the sequencing function and $|f|$ is the cardinality of elements that f sequences:

(8) $[\![\textit{respectively}]\!] = \lambda P_{\langle e,t \rangle} \lambda x \, . \bigwedge_{1 \leq i \leq |f|} f(P)(i)(f(x)(i))$

This uses f to reach into each member of the plural property P and applies it in turn to each member of the plural individual x, and ultimately turns (7) into something like $\textbf{ate}(\textbf{Alpo})(\textbf{Fido}) \wedge \textbf{ate}(\textbf{Whiskas})(\textbf{Felix})$. This represents what needs to be done, but suppresses inside f the details of how it's achieved.

---

2 Rewinding is an archaic mechanical process involved in the manipulation of certain forms of historical recorded media.

Answers are straightforwardly imaginable, of course, but they're not trivial. Perhaps the denotation of every linguistic expression includes some representation of words of which it's composed. Indeed, Potts (2007) proposes something along these lines to account for quotation and metalinguistic negation, and Morzycki (2011) considers (but rejects) the possibility for metalinguistic comparatives. It's worth noting, though, that it's not just the phonological string that's crucial but also its syntactic structure. All this would grant to a single adverb a considerable amount of power. If this were possible in general, we might expect language to be riddled with expressions that manipulate the relations between any two arbitrary discontinuous constituents in unrelated parts of the sentence, or even in preceding sentences. One might imagine an adverb, for example, that has the effect of swapping the two most recently mentioned verbs, or of predicating ugliness (for example) of the last member of the most recently mentioned conjoined DP. It's virtually a rejection of compositionality itself. All this, just for *respectively*?

I'm disinclined to go down this road, and presumably it's clear enough what road I find more appealing: *respectively* is a virus. It's not an ordinary adverb, and we shouldn't pretend that it is. It's weird. It's associated with the prestige form of the language, or in any case a formal elevated register. It's certainly not something kids say. It's not clear to me that it's something around which there are prescriptive injunctions or explicit instruction, but it's certainly something that's easy to botch. Indeed, pulling it off in extemporaneous speech with more than two- or three-member coordinate structures comes off as a flamboyant feat of linguistic acrobatics. Try it and you might find that your interlocutor applauds — or, more likely, that you've lost your train of thought.

The most interesting element of this, though, is that there's something distinctly self-conscious about it. It requires us to engage in a kind of metalinguistic reflection as we rewind the tape. This accords with Sobin and Lasnik's intuition that viruses involve real-time linguistic self-monitoring. It provides an independently-motivated understanding of the feeling of rewinding the tape. That's precisely the sort of thing that viruses manipulate.

This also presents what I'd like to suggest as another characteristic of viruses: they can be powerful, in a sense too powerful.[3] Certainly, *respectively* has powers far beyond the ken of ordinary adverbs. Treating it as a virus entails that it operates outside of the normal grammar. We therefore don't need to give the grammar the full power *respectively* demands, and thereby we avoid the prediction that affronts to compositionality like *respectively* should be commonplace. We needn't sacrifice compositionality itself on the alter of one grubby little adverb.

---

3 That may in some sense be implicit in the concept itself, and it seems to be a chord audible in Sobin's discussion.

Importantly, as far as I can see this isn't an argument against any particular account of the lexical semantics of *respectively*. It merely provides a theory of why its semantics might be special, a principled understanding of its weirdness. It also presents us with an interesting project: sorting out what analyses of semantic viruses have in common. To do that, it would help to briefly consider some more candidates.

4 "Zero" and negative numerals

Bylinina & Nouwen (to appear) provide an elegant analysis of the numeral *zero*, as in (9):

(9) a. Zero Americans enjoy natto.
     b. Most semanticists buy zero fancy cufflinks.

Their thesis is that *zero* is simply an ordinary numeral, no different from *three*, but that it predicates of a plurality that it has zero members. That would account for why it seems to be parallel to ordinary numerals in most respects, but it entails a bold ontological commitment: that there are pluralities with no members.

We are immediately in now-familiar territory: an uncomfortable and yet apparently necessary conclusion about the power of an apparently peripheral linguistic expression. Indeed, things might get worse. What to make of negative numbers?[4]

(10) a. ??It got warmer by negative ten degrees.
      b. ??Negative six people arrived.

It's relatively clear what (10a) is trying to mean: that it cooled by ten degrees. I'm not sure whether it quite manages it. I'm relatively sure that (10b) can't mean that six people left. But in both cases amid the wavering intuitions there's an impulse to throw up one's hands. God knows what sentences like these mean. They're nerdy half-jokes. Still, it would be nice not to have to regard them as simply ungrammatical either.

Yet if we regard *zero* and negative numerals as viruses, this is all expected. They're associated with a formal register. They're not often produced by toddlers. They require extensive tutoring to master. They are marked by a distinct sense of self-consciousness, a feeling of pushing at the boundaries of the language. We have to consciously reflect on what they mean, or what they're trying to mean. Our judgments are sometimes a little unsteady and wavering.

Most interesting, though, is again the question of power. Must we conclude that it's normal for pluralities to have zero members? If that were a possibility that's

---

4 This question was asked by Daniel Lassiter during a question period at SALT 27.

generally available, we'd expect languages to reflect it widely. We'd expect that other constructions would provide evidence for it. We'd expect it to occur across languages. After all, all things being equal we'd expect natural language ontology to be relatively similar across languages. Indeed, we'd expect to find manifestations of it in languages associated with cultures that have no concept of zero. If the possibility is straightforwardly available, it should be no harder to grasp than, say, a negated existentially-quantified proposition. All that holds true even more robustly for negative numbers. Treating these expressions as viruses addresses all these concerns.

As before, the analysis of these expressions would remain the same — but Bylinina & Nouwen's beautiful and alarming hypothesis is rendered less alarming, but no less beautiful.

This discussion raised two novel points. First, we've confronted the possibility that aspects of the ontology can themselves be viruses. Second, the notion of semantic viruses has helped us avoid an undesirable prediction not just about what's commonplace in a single language, but also what's commonplace across languages. Viruses allow us to avoid investing every language with the full power necessary in only some corners of only some languages.

5   Factor phrases

Precisely these points also arise in the domain of degrees. Factor or ratio phrases such those in (11) remain relatively little-studied (but see Gobeski 2009, in preparation, Gobeski & Morzycki to appear, Sassoon 2010a,b, Rett 2008):

(11)   a. Kyle's BMW is many times more expensive than his cufflinks.
       b. South College is half as tall as the new shiny new building.
       c. Fido has two times as much Alpo as Felix.

As Sassoon (2010a,b) points out, factor phrases provide a convincing argument for certain assumptions about the ontology of degrees. They demonstrate the scales of which degrees are members are *ratio scales*, the richest in the four-way taxonomy of measurement types. A ratio scale is, straightforwardly enough, one on which it's possible to compute ratios because it measures relative to a zero-point. That makes ratio scales richer — in the sense of having more structure — than mere interval scales, such as temperature, on which relative distances can be computed but ratios can't. And it makes them richer than ordinal scales, which don't support even that, and nominal scales, which are simply labels.

The interest of this is that it is incompatible with certain theories of what degrees are. On one popular view, a degree of height such as "six feet" can be understood as

the equivalence class of individuals that are six feet tall (Cresswell (1976) among others). This is a relatively simple and ontologically conservative way of thinking, but it doesn't yield a ratio scale. This issue doesn't arise with a sufficiently rich representations of degrees. On the assumption that degrees are a distinct atomic type in the model, they can be construed as elements of scales with a structure isomorphic to the real numbers (von Stechow 1984, Kennedy 1997 among others). Of course, that provides sufficient structure for ratios to be defined.

But across languages, this argument is harder to make. Most human languages were spoken by people with no explicit mathematical understanding of ratios. For this reason, factor phrases must be a relatively recent and relatively parochial development. Many languages lack them, and so lack (at least this piece of) evidence for ratio scales. Worse, even differential measure phrases like *three feet taller*, aren't universal, so many languages lack evidence for even interval scales. And, following this line of thought to its logical conclusion, in many languages, there is little evidence for degrees at all (Bochnak 2013).

That leaves us with the conclusion that the ontology underlying natural languages is not only not universal, but that it can vary dramatically in its complexity. That runs afoul of the appealing intuition that the basic structure of the ontology should be relatively invariant because it has to interface with the same language faculty across languages.

There is, however, a way to resolve this tension. Perhaps factor phrases are viruses? This may be stretching the point, but it's certainly the case that they're not typical of child language and they require explicit instruction. Arguably, they're associated with formal language, setting aside uses that don't involve actual ratios like *a million times as depressed*. Moreover, they have clearly have the two additional properties of semantic viruses I've pointed to above. In a clearly definable sense, they require more power in the semantics — more structure — than we otherwise have reason to assume. And they are absent in many languages, and therefore unlikely to be a feature of the language faculty itself.

Indeed, if one is in a sufficiently radical mood, one might consider measure phrases themselves in a similar spirit, or even degrees themselves. It might be best to get off this train considerably short of that station, but it's headed in an interesting direction. Lasnik & Sobin (2000) regard the viral *whom* as kind of historical relic. Perhaps viruses can play the opposite role too, as mechanisms by which complexity enters a language before becoming fully "nativized" into the grammar.

6 A final remark

I've suggested that the semantics, like the syntax, may be infected with grammatical viruses — phenomena that are weird in a precise sense that needs to be recognized

in linguistic theory. In semantics, it's helpful to attend to two additional properties of viruses: first, that they require more power than we'd like to make available to the grammar more broadly, and second, that they are unusual not only within a language but crosslinguistically. Much rests on how to construe "more power", of course, but "more structure" seems a reasonable first approximation. This raises the possibility of regarding certain elements of the ontology or certain aspects of its structure as viruses.

If the case for viruses in the semantics is at all convincing, we might seek them in the phonology too. The notion of "crazy rules" (Bach & Harms 1972) might provide a starting point (Anne-Michelle Tessier, p.c.). And if that case is convincing, it's worth sorting out in detail what analyses of semantic viruses have in common, and, ultimately, whether it's possible to develop a general theoretical framework for their analysis.

References

Bach, Emmon & Robert T. Harms. 1972. How do languages get crazy rules? In Robert Stockwell & Ronald Macaulay (eds.), *Linguistic change and generative theory*, 1–21. Bloomington: Indiana University Press.

Bochnak, M. Ryan. 2013. *Cross-linguistic variation in the semantics of comparatives*: University of Chicago dissertation.

Bylinina, Lisa & Rick Nouwen. to appear. On 'zero' and semantic plurality. In Dan Cameron Burgdorf & Jacob Collard (eds.), *Proceedings of proceedings of semantics and linguistic theory (SALT) 27*, eLanguage.

Cresswell, Max J. 1976. The semantics of degree. In Barbara H. Partee (ed.), *Montague grammar*, 261–292. New York: Academic Press.

Dalrymple, Mary & Andrew Kehler. 1995. On the constraints imposed by respectively. *Linguistic Inquiry* 26(3). 531–536.

Eggert, R. 2000. Grammaticality and context with respect to and... and or... respectively. In *Proceedings of the chicago linguistics society (CLS) 36*, Chicago: Chicago Linguistics Society.

Emonds, Joseph E. 1986. Grammatically deviant prestige constructions. In Michael Brame, Heles Contreras & Frederick Newmeyer (eds.), *Festschriftfor Sol Saporta*, 93–129. Seattle, Washington: Noit Amrofer.

Farkas, Haley. 2015. The semantics of *Vice Versa*. B.A. senior thesis, Michigan State University.

Gawron, Jean Mark & Andrew Kehler. 2004. The semantics of respective readings, conjunction, and filler-gap dependencies. *Linguistics and Philosophy* 27(2). 169–207.

Gobeski, Adam. 2009. *'Twice' versus 'two times' in phrases of comparison*: Michigan State University MA thesis.

Gobeski, Adam. in preparation. *Factor phrases*: Michigan State University dissertation.

Gobeski, Adam & Marcin Morzycki. to appear. Percentages, relational degrees, and degree constructions. In Dan Cameron Burgdorf & Jacob Collard (eds.), *Proceedings of proceedings of semantics and linguistic theory (SALT) 27*, eLanguage.

Kac, Michael B. 1987. Surface transitivity, respectively coordination and context-freeness. *Natural Language and Linguistic Theory* 5(3). 441–452.

Kay, P. 1989. Contextual operators: respective, respectively, and vice versa. In *Proceedings of the berkeley linguistics society (BLS) 15*, 181–192.

Kennedy, Christopher. 1997. *Projecting the adjective: The syntax and semantics of gradability and comparison*: UC Santa Cruz dissertation. Published in 1999 by Garland, New York.

Lasnik, Howard & Nicholas Sobin. 2000. The who/whom puzzle: On the preservation of an archaic feature. *Natural Language and Linguistic Theory* 2. 343–371.

McCawley, James D. 1968. The role of semantics in a grammar. In Emmon Bach & Robert Harms (eds.), *Universals in linguistic theory*, 124–169. New York: Holt, Rinehart and Winston.

McCawley, James D. 1998. *The syntactic phenomena of English*. Chicago: University of Chicago Press. Second edition.

Moltmann, Friederike. 1992. *Coordination and comparatives*: MIT dissertation.

Morzycki, Marcin. 2008. Nonrestrictive modifiers in nonparenthetical positions. In Louise McNally & Christopher Kennedy (eds.), *Adjectives and adverbs: Syntax, semantics, and discourse* Studies in Theoretical Linguistics, 101–122. Oxford: Oxford University Press.

Morzycki, Marcin. 2011. Metalinguistic comparison in an alternative semantics for imprecision. *Natural Language Semantics* 19(1). 39–86.

Okada, S. 1999. On the function and distribution of the modifiers respective and respectively. *Linguistics* 37(5). 871–903.

Potts, Christopher. 2007. The dimensions of quotation. In Chris Barker & Pauline Jacobson (eds.), *Direct compositionality*, vol. 14 Oxford Studies in Theoretical Linguistics, 405–431. Oxford: Oxford University Press.

Pullum, Geoffrey K. & Gerald Gazdar. 1982. Natural languages and context-free languages. *Linguistics and Philosophy* 4(4). 471–504.

Rett, Jessica. 2008. *Degree modification in natural language*: Rutgers University dissertation.

Sassoon, Galit. 2010a. The degree functions of negative adjectives. *Natural Language Semantics* 18(2). 141–181.

Sassoon, Galit. 2010b. Measurement theory in linguistics. *Synthese* 174(1). 151–180.

Schütze, Carson. 1999. English expletive constructions are not infected. *Linguistic Inquiry* 30(3). 467–484.

Sobin, Nicholas. 1994. An acceptable ungrammatical construction. In Susan D. Lima, Roberta L. Corrigan & Gregory K. Iverson (eds.), *The reality of linguistic rules*, 51–65. Amsterdam: John Benjamins Publishing.

Sobin, Nicholas. 1997. Agreement, default rules, and grammatical viruses. *Linguistic Inquiry* 28(2). 318–343.

von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3. 1–77.

Stockwell, Robert, Paul Schachter & Barbara Partee. 1973. *The major syntactic structures of English*. New York: Holt, Rinehart, and Winston.