Person Asymmetries in English Controlled and Exempt Anaphors

Dennis Ryan Storoshenko
University of Calgary

Presented at the Manitoba Workshop on Person
Sept 22-23, 2017
Winnipeg, MB
Genesis of the Research

Betsy: “You should submit something”

Me, thinking: But the term just ended, and I have time to finally figure out how to use this Twitter tool I’ve been meaning to play with

So, this started as a simple project to replicate a person asymmetry I found in part of my thesis work, but never really did anything with...

Fairly basic observation: exempt anaphors in English (Condition A violations) tend to largely be first or second person

(1) On behalf of myself and everyone at Air Canada...
Why this is Interesting

Since I did that corpus work on English reflexives in 2008, a new analysis for *self* pronouns has appeared:

Kratzer (2009) argues that reflexives are built from the interaction between a featureless variable, a binder at $\nu$, and the $\phi$-features of the antecedent

The resulting feature bundle gets its final form after spellout

\[ \text{(2) Nomi} \ [\nu \lambda n \ [\text{hurt} \ n]] \quad \rightarrow \quad \text{“Nomi hurt herself.”} \]

The consequence of this is that Condition A effects are epiphenomenal; violations should be impossible to spell out
But...

We know Condition A violations occur in natural speech (Storoshenko 2008 for numbers from Treebank 3)

Two possible explanations:

1. Those exempt anaphors are just inserted whole cloth, and there’s no feature checking story to tell

2. We can invoke some sort of covert antecedent to explain the source of the features

I’m going to be pushing on explanation 2 here
Gameplan

Go through the corpus methods and findings

Sidetrack into a discussion of red herrings that tell us something about control

Building up the idea that PRO seems to be bindable by speech act participants

Sketching an account of how those same participants can be the binders of our apparently exempt anaphors
Prior Data

In 1 millions words each of written (WSJ) and spoken (Switchboard) text, 22 exempt anaphors, only one of those from written

The spoken ones were all first or second person

No real attention paid to them beyond speculating they might be some sort of politeness marker

With a total 1071 reflexives (singular and plural), this is 0.021% of the usage
Replication via Twitter

Searched the entire *self* paradigm, set for a 100 mile radius around Calgary (Gentry 2015, R-Core Team 2017)

This gathers roughly one week’s worth of tweets, about June 2-9, 2017

Ran a program to do some cleanup on the data (remove retweets, duplicates, other bugs)

Manually annotated for function of the *self* form

  Argument (including inside PPs), Adverbial (Inclusive and Exclusive), Adnominal, Exempt

Coarser grained than we might like, but the data is all still there
The Numbers

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Exempt</th>
</tr>
</thead>
<tbody>
<tr>
<td>myself</td>
<td>766</td>
<td>73</td>
</tr>
<tr>
<td>yourself</td>
<td>943</td>
<td>19</td>
</tr>
<tr>
<td>himself</td>
<td>281</td>
<td>1</td>
</tr>
<tr>
<td>herself</td>
<td>85</td>
<td>1</td>
</tr>
<tr>
<td>itself</td>
<td>168</td>
<td>0</td>
</tr>
<tr>
<td>oneself</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>

That gender disparity is interesting. Just out of curiosity, I re-did *himself* and *herself* using Toronto, and got a similar imbalance.

Same basic result for the exempt cases, almost all 1st or 2nd person.

Without a word count for all of Calgary twitter over that timeframe, it’s impossible to make any claims about % of total words, but we are up to 0.042% of the singulars.
The Red Herrings

There are many examples like this:

(5) Or himself in a mirror (@masson_yves)

Conversational sentence fragment; once we resurrect the rest of the sentence, this would be fine (so called this an argument)

(6) What Trump its doing is feeding his base with B.S saying himself and guys [sic] supporters are under siege. (@Angelo_Calgary)

The antecedent DP *Trump* is quite far away from the *himself*, but we can squabble over what the subject of *saying* is meant to be (Trump or the BS?)

With a proliferation of relative clauses and control, not to mention fragments, there are a lot of at-first-glance antecedentless *self* forms to filter out
Control and Reflexives

For the purposes of this talk, I’m going to assume PRO is actually there in the syntax and semantics (i.e. the infinitive clause is a fully saturated predicate)

(7) Wolfgang wants PRO to outdo himself.

In the Kratzer story, PRO would be the one binding the variable that eventually spells out as himself. Those φ features must come from somewhere though.

Now we just have to invoke our favourite story about how the subject binds the PRO.

Of a piece with Kratzer’s story for reflexives is Borer’s (1989) account of control being achieved by the binding and transmission of features via AGR

There’s an issue of counter-cyclicity here, assuming we spell out in phases, that I am going to acknowledge, but fail to solve.
Controlling Sartorial Anxiety

Another troublesome example:

(8) I’m a grown man who doesn’t know how to dress myself properly. (@n3rdbassist)

One would imagine the PRO on *dress* is controlled by the subject of the *know* clause, which is bound by *grown man*. We can easily make this work:

(9) I’m a grown man who doesn’t know how to dress himself properly.

This one fairly clearly predicates of the first person subject that he is a sartorially-challenged grown man. It’s just a very complexly-built property.

We could just say the usual thing about the *myself* being an exempt anaphor for emphasis, but there’s a problem with that...
Limited Exemption

Let’s take a more canonical exempt case from the twitter dataset:

(10) I blame the Chemtrails. Only advanced physical specimens like yourself can resist their effects. (@solarwolfman)

We can replace this *yourself* with plain old *you*, and still get a grammatical sentence with the same basic truth conditions.

The same goes for all of the *Lucie and myself* examples from Reinhart and Reuland (1993)

(11) *I’m a grown man, who doesn’t know how to dress me/him, properly.*

This position cannot be similarly replaced, so the reflexive is apparently obligatory.
Multiple Controllers?

But it’s not the case that this is some sort of obligatorily reflexive predicate:

(12) I’m a grown man who doesn’t know how to dress his kids/his dogs/a salad properly.

The *myself* is also sloppy under ellipsis; it’s fairly clearly bound, with no strict reading:

(13) I’m a grown man who doesn’t know how to dress myself properly, and Felix is too.

What seems to be happening is that when *dress* is a reflexive predicate, we have Kratzer’s featureless variable, getting two different feature sets, either first person or third

But, this need not be happening directly; what we may be seeing is two different binders for the PRO
Another example to get at this idea:

(14) Any ideas on how to clone myself? (@MavenBkkg)

This can just as easily have a second person object, but not third:

(15) Any ideas on how to clone yourself?
(16) *Any ideas on how to clone herself/himself/itself?

Generously, we can imagine the missing matrix clause here was *Can you give me*, even though it reads more like *Do you have*

It sounds clunky, but this third person is ok:

(17) Any ideas on how to clone oneself?

The usual story for arbitrary PRO is that there’s some arbitrary binder way up high.
Solo Dining

Yet another:

(18) Is it weird to take yourself out to dinner? (@MeghanBurton)

Again, we can just as easily have first person, but not third (except oneself)

(19) Is it weird to take myself out to dinner?
(20) *Is it weird to take herself/himself/itself out to dinner?

If this is ellipsis, then what we are eliding is complementizer for and the subject of the infinitival clause (meaning this isn’t even really control)

But, that should mean that (20) is equally possible, and my instinct here is that it is at least somewhat degraded
Still Unconvinced?

To my ear, the answers are sharper:

(21) No, it’s not weird to take myself out to dinner!
(22) No, it’s not weird to take yourself out to dinner!
(23) *No, it’s not weird to take herself/himself/itself out to dinner!

Another test to try out is with plurals. If we really are eliding the for and a subject, we should also be able to imagine a plural subject licensing a reciprocal or plural reflexive:

(24) Is it weird to take each other out for dinner?
(25) *Is it weird to take themselves out for dinner?

(24) is perfectly natural if the each other is speaker and addressee, much harder to imagine having elided Is it weird for Wolfgang and Kala to take themselves/each other out for dinner?
Speech Participant Binding

As we already saw with the discussion of imposters (Collins and Postal 2008), AUTHOR and ADDRESSEE are available high in the tree.

Aside from being a source for co-reference, I am proposing that these can bind PRO, leading to paradigms like (18)-(20) where 1st and 2nd reflexives under control were fine, but 3rd was out.

This gets us to a point of accepting that control need not always take the closest possible binder, (though AUTHOR was not necessary in the *dress myself* example)

Now, if we can accept the Participants as binders for PRO, we can think about what mileage we might get out of extending this to the exempt anaphors.
Denotations for Pronouns

Returning to our constructed non-exempt sentence from earlier:

(26) Only advanced physical specimens like you can resist their effects.

We can assume a fairly standard Kaplan (1989)-inspired denotation for the first and second person pronouns:

(27) \[ [I]^{c,i,g} = \text{AUTH}(c) \]
(28) \[ [\text{you}]^{c,i,g} = \text{ADDR}(c) \]

The pronoun in (26) has its reference determined relative to the context, which provides the addressee. In other words, co-reference.
The Speech Act Projection

This is where I take a dive into naive work in progress, more reading to do, so bear with me.

Up at the top, there is a three-place predicate, taking the propositional content of the sentence, type \( \langle s,t \rangle \), and two type \( \langle e \rangle \) arguments for the AUTHOR and ADDRESSEE:

\[ \lambda p. \lambda x. \lambda y. \lambda s. \exists e. [\text{SPEECH-ACT}(e,s) \land \text{content}(e,p) \land \text{Addressee}(e,x) \land \text{Author}(e,y)] \]

Suppose we could leave one of those minimal pronouns unbound in TP, and get the binder up this high:

Just as an external argument saturates the transitive predicate and binds \( n \) to get a reflexive, the DP destined to be AUTH does the same (admittedly over much longer distance).
The Speech Act Projection

This is where I take a dive into naive work in progress, more reading to do, so bear with me.

Up at the top, there is a three-place predicate, taking the propositional content of the sentence, type \( \langle s,t \rangle \), and two type \( \langle e \rangle \) arguments for the AUTHOR and ADDRESSEE.

\[
\lambda p \lambda x \lambda y \lambda s. \exists e [ \text{SPEECH-ACT}(e,s) \land \text{content}(e,p) \land \text{Addressee}(e,x) \land \text{Author}(e,y)]
\]

Suppose we could leave one of those minimal pronouns unbound in TP, and get the binder up this high:

Just as an external argument saturates the transitive predicate and binds \( n \) to get a reflexive, the DP destined to be AUTH does the same (admittedly over much longer distance).
Exemption as Participant Binding

Back to the example:

(29) I blame the chemtrails. Only advanced physical specimens like yourself can resist their effects. (@solarwolfman)

We would also have to assume that binding from the speech act projection also provides a feature to spell out as a *self* form.

Assimilating the adverbial cases to this analysis would require a similar syncretism.

But instead of co-reference with AUTHOR, the exempt anaphor is bound directly by it, in a sense anchoring the anaphor to the speech act: “you who are definitely my current interlocutor with no way for context to mess with that”
Exempt anaphors in English are almost exclusively first or second person.

Control allows multiple possible binders for PRO.

These can include AUTH and ADDR (and maybe the USH can solve cyclicity?, thanks Martina and Betsy).

We can bring the exempt *self*-pronouns into the realm of binding if we allow AUTH and ADDR to be binders over seemingly arbitrarily long distance, explaining the person asymmetry.

As a side benefit, we get a flicker of a meaning difference between using a regular pronoun versus the exempt *self*-pronoun.
Acknowledgements & References

Thanks first of all to the organizers for putting on this workshop. Thanks also to all of you for sticking around until the end, the tweeters of Calgary, and colleagues back at the U of C for putting up with my pestering about judgements on those control cases for the past few weeks.


Gentry, Jeff. 2015. twitteR: R-based Twitter client. URL http://lists.hexdump.org/listinfo.cgi/twitter-users-hexdump.org


Another piece of this is what I was talking about back at the CLA

In that 2008 corpus study, I had also found that only about half of all English *self* pronouns are even in “canonical” argument positions

I reported some felicity judgements establishing that at least some adverbial uses of the *self* pronouns can be assimilated to the Kratzerian analysis

These would be the hardest ones to fit into that analysis
The 3rd Person Examples

(3) Only one person to blame for this & that’s herself (@Putter99)

This is a predicative use, and we can easily say something about the *that* being anaphoric to the person (Kathy Griffin)

(4) The himself who leaves the scene after crashing his car and injuring pregnant women? (@DannyAustin_9)

Following the conversation thread back, there had been a discussion of a news headline using *himself*, so this seems to actually be a whole cloth referential use

Still, there is that person asymmetry. And this task proved to be harder with twitter data than I remember it being with Treebank