Features, Identity, and ‘Yourself’*

Byron Ahn (bta@princeton.edu)
Princeton University
NELS 49, Cornell University

1 Introduction

- English reflexive nominals are composed of two nominals: (possessive) pronominal + self nominal

(1) We will not embarrass our selves

(2) **GENERALIZATION ON ϕ-MATCHING**

In English reflexives, the nominals ϕ-match the antecedent of binding

- This is introduced in textbooks and presupposed by researchers across a spectrum of analyses (cf. discussion in Sundaresan 2018)
  - Kratzer (2009), on her derivational analysis of matching ϕ-features: “We don’t build [nonagreeing reflexives] to begin with.” (p.196)
  - Hicks (2009), who does not adopt a derivational analysis of matching ϕ-features: “[...]anaphors and their antecedents do share the same values for [ϕ].” (pp.107-108)

**TWO QUESTIONS**

- **Empirical Question**: How valid is this generalization?
- **Theoretical Question**: How do we derive this generalization (to the extent it is valid)?

- Some findings, to be motivated:
  - ϕ-features in a self-phrase can mismatch its antecedent, undercutting accounts of binding that directly invoke ϕ-AGREE (or similar syntactic operations) as a core component of binding
    - i.e., reflexivity is interpreted separately from the ϕ-features – not surprising given that many languages lack any manifestation of ϕ-features in reflexives
  - English 3.sg reflexive pronouns behave distinctly
    - Unlike other bound pronouns, they seem to require a derivational analysis of ϕ-features
    - i.e., There are two derivational routes to ϕ-features
  - “Binding” operations/constraints are distributed across the Grammar

*This work has largely been influenced by collaboration with Laura Kalin, who I owe a great debt of gratitude. Additional thanks are owed to the audience and participants of the Ana-Log Workshop at Harvard this fall, where some of this work was also presented.
2 Internal Structure of Reflexive Self-Phrases

- English reflexives are morphologically complex (see also Postal 1966:182)
  - Possessive pronoun (myself) + a head noun (myself)
  - (This is typologically common: the reflexive anaphor is composed of a possessor and a inalienably possessed nominal)

There is a self component, which is a nominal:
- Consider the object-incorporation-like compounding process:
  
  (3) a. They are completely self-reliant.
      b. They rely on themselves completely.
  
  (4) a. He is a self-described polyglot.
      b. He describes himself as a polyglot.
  
  (5) a. We chose the self-install option.
      b. We chose the option to install it ourselves.
  
  (6) a. This is a self-driving car.
      b. This is a car that drives itself.

- What is is adjacent to the V is a bare N, as with other Ns:
  
  (7) a. She is an Eagles-supporting Philadelphian.
      b. She is a Philadelphian who supports the Eagles.

**HISTORICAL CHANGE**

- In older forms of English (and still in other Germanic languages), the self morpheme is an adnominal intensifier\(^1\)
  - It lacked nominal distribution, didn’t inflect like a nominal, etc. (cf. Keenan 2002)
- Now self is certainly nominal
  - Given evidence like compounding, pluralization, etc.

What precedes self is a possessive pronoun:
- Consider the form of the pronoun:
  
  (8) a. I will defend myself.
      b. You can do it yourselves.

- Headlinese allows for null bound pronoun possessors
  
  (9) a. ...Bill O’Reilly embarrasses self, colleagues, country... (http://wapo.st/1TYFgoH)
      b. Bill O’Reilly embarrassed himself, his colleagues, and his country.

\(^1\)In fact, this difference is almost certainly related to the other differences between English reflexive anaphors and reflexive anaphors in other Germanic languages (e.g., absence/presence of possessive pronouns within the anaphor, and possibly even (un)availability of long distance binding).
In addition, there is some **invisible reflexive morphosyntactic glue**

- It is detectable through *allomorphy* in 3rd person contexts (for certain varieties of English)

  (10) a. They did it *them/* *their* selves.

  b. After spending two years in meditation, *their/* *them* selves were fully realized.

  - A reflexive morpheme, $\text{REFL}^0$, triggers 3rd person pronouns to surface as ‘accusative’ in certain structure-based contexts
  - What disturbs the locality shows that the $\text{REFL}$ morpheme is distinct from the $\sqrt{\text{SELF}}$ morpheme and the $\text{D}$ morpheme

- It is also detectable through *allosemy*: $\sqrt{\text{SELF}}$ lacks clear content in reflexive contexts

  - It gets a reified substantive reading (cf. Safir 1996) in the absence of $\text{REFL}^0$
  - But $\text{REFL}^0$ triggers a lack of semantic contribution by the root $\sqrt{\text{SELF}}$ (cf. camouflage constructions; Collins et al. 2008, Collins and Postal 2012)

- Ahn and Kalin (*forthcoming*): A reflexive morpheme that sits between $\text{D}$ and $\text{N}^2$

(11) *yourself* (reflexive)

\[\begin{array}{c}
\text{DP} \\
\text{DP} \quad \text{D}' \\
\text{your} \quad \text{D} \\
\text{REFLP} \\
\text{REFL} \quad \text{NP} \\
\text{DP} \quad \text{N}' \\
\text{your} \quad \text{N} \\
\text{self}
\end{array}\]

### 3 Features in the Self-Phrase

- In this structure of reflexive self-phrases that we have motivated, there are two nominals (pronominal possessor + $\sqrt{\text{SELF}}$) inside of the anaphor

  - What gives rise to the nominal features (“$\phi$-features”) that manifest on these two nominals?

    - The pronoun (person, number, gender, animacy, genericity)
    - The $\sqrt{\text{SELF}}$ nominal (number)

- How do we derive the $\phi$-Matching generalization in (2), to the extent that it is correct?

---

2One idea might be that *own* manifests this $\text{REFL}$ head; this cannot be the case, as *own* does not entail reflexivity: ‘*For my friends, their true selves were being revealed to them, but my own self never revealed itself to me*’. In a case like this, where *own* occurs between *my* and *self*, you are not guaranteed a reflexive interpretation.

3This raises interesting issues with a Reinhart and Reuland 1993 style typology of anaphors as $\pm R$ and $\pm \sqrt{\text{SELF}}$. Under this analysis of English, English anaphors might be seen as $\pm R$ and $\pm \sqrt{\text{SELF}}$, in the way that Anagnostopoulou and Everaert (1999) describe for Greek. There are issues to work out if this is how English anaphors are to be analyzed (in the same way that issues arise for Greek). (*Thanks to Nikos Angelopoulos for pointing this out to me.*)

4I set aside the case features on the pronoun, which some may want to include alongside these other nominal features. For a discussion of that, see Ahn and Kalin (*forthcoming*). The case features of the whole self-phrase are set aside as an open question.
3.1 Two Analytical Approaches to $\varphi$-Features

- Some approaches treat bound pronouns as $\varphi$-deficient “minimal pronouns” that get their $\varphi$-features value during the derivation – “Derivational Approaches”
  - Kratzer: a functional head\(^5\) values reflexive anaphors’ $\varphi$-features via Feature Transmission
  - Others: AGREE with the nominal antecedent values reflexive anaphors’ $\varphi$-features
    - (e.g., Heinat 2006, Reuland 2006, Rooryck and Vanden Wyngaerd 2011)
- Others treat bound pronouns as having their $\varphi$-features specified as soon as $\varphi$-bundles are merged/built
  - This sort of idea is compatible with other works that do not for syntactic feature matching, but rather rely on interpretive compatibility (e.g., Heim 2008, Hicks 2009, Safir 2014)
  - Heim argues that $\varphi$-features correspond to presuppositions in the semantics, effectively (though not directly) causing $\varphi$-features of a bound pronoun to match those of the antecedent

<table>
<thead>
<tr>
<th>THEORETICAL QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>• What is/are the mechanism(s) that yield feature matching?</td>
</tr>
<tr>
<td>• Are the pronouns in reflexive self-phrases uniform, regarding whether their $\varphi$-features are derivationally checked/valued?</td>
</tr>
</tbody>
</table>

3.2 English Mismatches

- Let us continue by denying\(^6\) the universality of the generalization that the $\varphi$-features in a reflexive must match the $\varphi$-features of the antecedent
  - The pronominal possessor and the antecedent of binding need not match in $\varphi$-features:\(^7\)

\[
\text{(12) a. } I \text{ wouldn’t blame your self!} \\
\text{[#:sg, } \pi:1, [\pi:2] \text{]}
\]

\[
\text{b. } \text{Everyone loves them selves} \\
\text{[#:sg, } \pi:3, [#:pl, \pi:3] \text{]}
\]

\[
\text{c. } \text{Each of us is defending our selves} \\
\text{[#:sg, } \pi:3, [#:pl, \pi:1] \text{]}
\]

\[
\text{d. } \text{At least one of you believes in your self} \\
\text{[#:sg, } \pi:3, [\pi:2] \text{]}
\]

\[
\text{e. } \text{Your majesty needs to protect your self} \quad \text{(Collins and Postal 2012)} \\
\text{[#:sg, } \pi:3, [\pi:2] \text{]}
\]

---

\(^5\)This head, in turn, will have gotten its features under a relationship with a separate nominal; e.g., the subject.

\(^6\)The generalization can be made universal if (i) the feature-matching syntactic operation is more nuanced in how/when it applies, and/or (ii) the syntactic $\varphi$-features that value the features of the pronoun need not be the overtly manifested $\varphi$-features on the antecedent.

\(^7\)Data like (12c) show that Kratzer’s (2009) analysis of German bound pronouns does not obviously extend to the pronouns in English reflexives; the form of be (‘is’) suggests that $V^0$ has 3sg $\varphi$-features, which should be incompatible with building our from a minimal pronoun under Feature Transmission.
There can be mismatch between the √SELF and the antecedent:

(13) a. You guys pushed your self, drove your self, sacrificed, trained and competed
       [#\text{pl}]  [#\text{sg}]  [#\text{sg}]  
       (M.Romney 2002)

       b. The team credits them selves
       [#\text{sg}, \pi:3]  [#\text{pl}]

Previous work has been established that features of the maximal DP (whose head N would be √SELF in English) can differ from the features of the pronoun (cf. Greek clitic doubling; Iatridou 1988, Anagnostopoulou and Everaert 1999)

There can be mismatch between the √SELF and the pronoun:

(14) a. \% Everyone loves them self
       [#\text{pl}, \pi:3]  [#\text{sg}]

       b. \% We each did it our self
       [#\text{pl}, \pi:1]  [#\text{sg}]

       c. \% We all need to ask our self [a very serious question]
       [#\text{pl}, \pi:1]  [#\text{sg}]

       (ABC Nightline)

⇒ These are serious problems for analyses that treat the generalization in (2) as a premise/explanandum

At the same time, not just anything goes


       b. * He behaves your self.

⇒ Whatever rules out these forms must not make any appeal to a mechanism that requires a matching of φ-features

3.3 Referential Construal

Intuitively, the constraint seems to be interpretive – the pronoun and anaphor must be interpretively compatible

(similar views in, e.g., Heim 2008, Hicks 2009, Safir 2014)

Interpretive compatibility for bound pronouns appears to be something like ‘weak identity’

each of us and ourselves are interpretively identical enough

Similar to proxy readings: statues and a person are identical enough

(16) Ringo admired himself (=the statue of Ringo)

‘him’ is identical enough to ‘Ringo’

It is not identity between syntactic features that is necessary in English

‘each of us’ (3sg) can be construed as identical to ‘us’ (1pl)

This weak identity is enough to rule out data like (15)

‘he’ cannot be construed as identical to ‘you’

\^See Reuland and Winter 2009 for a formal analysis that derives this ‘identical enough’ quality for proxies.
What matters is how the pronoun/antecedent is construed – implicating the role of the interpretation of (not the formal features of) the antecedent

Discussed by Collins and Postal (2012) in their investigation of imposters

17. [Spoken to a king]
   Your majesty must protect yourself/himself.

18. [Spoken by a parent to a child]
   Mommy and Daddy need some time to ourselves/themselves.

They note the generalization that the features of the anaphor can match the features of the “ultimate antecedent” (in the case of addressing ‘your majesty’, the ultimate antecedent can be construed as 2nd person or 3rd person masculine) – Collins and Postal 2012:Ch.14

‘Each of us’ can be construed as ‘the group including the speaker’, allowing felicitous binding of ‘ourselves’ – note that context may drive the gender choice in the matching condition

19. [Spoken by a single woman in a group of women]
   Each of us is doing it ourselves/herself.

20. [Spoken by a single man in a group of men]
   Each of us is behaving ourselves/himself.

‘They’ can be construed as an indefinite9 individual of unknown/irrelevant gender, allowing felicitous binding by QPs

21. [Spoken about a group of mixed-gender artists]
   Every artist ought to express themselves/ourselves/himself.

For some speakers, ‘they’ can even occur where the gender identities of the members of the QP are the same and known to the speaker10

22. [Spoken about an individual of unknown gender]
   Whoever that is ought to control themselves/himself.

23. [Spoken about an individual with non-binary gender identity]
   Kim wrote a book by themselves/himself.

(lots of inter-speaker variation is observed in contexts like these)

Groups of individuals can be construed as a collection of individuals or a single inanimate entity

24. The local football team organizes the weekly tailgate itself/themselves.

φ-features of the relevant type may not even be syntactically represented on the antecedent

Common nouns in English (even ones that appear gendered, according to social norms) don’t have gender φ-features (examples from Ackerman 2018)

25. a. At the farmhouse, the cowgirl left his lasso in the kitchen.
   b. At the Halloween party, the cowgirl left his lasso in the kitchen.

“The feminine definition associated with cowgirl is thus defeasible, since gender agree-

9There are large groups of people (including me) for whom ‘they’ can also be used in contexts with definite individuals, where the speaker knows the referent to have male/female gender identity (and has possibly already committed to it in conversation), but does not invoke it (again). e.g., ‘I know {the secret winner of the contest} has told {his/her} friends that {he/she} won. [...] {Their} opponents have not yet been informed.’ (See also Conrod 2017.)

10This also does not address non-binary gender uses of they/them. See Ackerman 2018 for discussion and references.
Antecedent nominals like the cowgirl cannot serve to provide a $\varphi$-value of [$\gamma$: +$\mathbf{f}$] to any reflexive pronoun with an unvalued gender feature.

- Names might have gender features (Bjorkman 2017), but their gender features must be flexible enough to account for gender identity ("social gender", Ackerman 2018).
  - Thus expressions like ‘Sue likes himself’ are not ungrammatical; where deviant, it is infelicitous (Conrod 2017, 2018)
    - Sue ⇒ Individual of female gender ⇒ infelicitous as binding himself
    - Sue ⇒ Individual of non-binary gender ⇒ infelicitous as binding herself
  - A lack of gender $\varphi$-features on the antecedent is problematic for analyses where bound pronouns are $\varphi$-deficient and rely on valuation from the antecedent.
- Instead what constrains felicity of particular bound pronouns in contexts like (25) is the interpretation of $\varphi$-features in context.
  - What matters is construal\(^\text{11}\) as weakly identical:

\[
\text{(26) WEAK IDENTITY CONDITION IN ENGLISH}
\]
The bound pronoun in an English reflexive must be able to be construed as (weakly) interpretive identical to its antecedent.

- More work is necessary to define exactly how this notion of weak identity is precisely defined, constrained, and implemented by the grammar.
  - Indeed, this is a (purposefully) weak theory, which may be able to generate expressions that are deemed unacceptable.
  - Weak Identity alone is not enough; auxiliary constraints (linguistic and social) will be necessary to capture the range of (un)acceptable data.
- To expose the need for auxiliary constraints, consider a gap in the mismatches that we have seen.
  - Not attested: 3.SG anaphors with a 1st/2nd person or PL antecedent
    - (27) a. The football team organizes the weekly tailgate itself/themselves.
      b. The football players organize the weekly tailgate *itself/themselves.
    - (28) [Spoken by a single man in a group of men]
      a. Each of us is behaving himself/ourselves/%ourself.
      b. We are each behaving *himself/ourselves/%ourself.
- This suggests that there is a morphosyntactic division in these reflexive pronouns, which goes beyond the interpretive constraint in (26)
  - We will return to this.
- Let us turn to contexts that provide weak identity but which do not allow all types of mismatch.
  - Reinforcing that there are additional constraints beyond the condition on weak identity supplied by anaphor binding.

\(^{11}\)I leave open what grammatical/non-grammatical variables/operations should be used in modeling “construal”.
3.4 Counter-Indexical Contexts

- One reason to believe construal as weakly identical is what matters in English are so-called counter-indexical (CID) contexts such as (29):

  (29) [Speaker A is going to the airport shortly, and asks Speaker B whether it’s a good idea to bring food or buy food on the plane. B replies...]
  a. If I were you, I’d do myself a favor and bring food! [✓1 ant, 1 ana]
  b. If I were you, I’d do yourself a favor and bring food! [✓1 ant, 2 ana]
  - In this sort of shifted context\(^\text{12}\) that can be introduced with ‘if I were someone else’, the local syntactic antecedent of binding is a 1st person pronoun, but the self-phrase can contain a 2nd person pronoun
  - Because they are construed as identical in the CID context where ‘I’ = ‘you’

- As a first constraint, consider the fact that this CID context only allows for $\varphi$-mismatch with modals

  (30) [same context]
  a. When I was you (in a dream), I did myself a favor and brought food! [✓1 ant, 1 ana]
  b. *When I was you (in a dream), I did your self a favor and brought food! [✓1 ant, 2 ana]
  - Which $\varphi$-features a pronoun in a reflexive self-phrase can bear —i.e., the availability of $\varphi$-mismatch— depends on irrealis mood

- $\varphi$-mismatching in these CID contexts is constrained in complex ways
  - The reflexive pronoun cannot be 3.SG: compare (29)–(32) to (33)

  (31) [Speaker A looks on at Speakers B and C, who are about to move the fridge, just the two of them. A says...]
  a. If I were you guys, I wouldn’t try to move the fridge by myself! [✓1 ant, 1 ana]
  b. If I were you guys, I wouldn’t try to move the fridge by yourselves! [✓1 ant, 2.pl ana]

  (32) [Speakers A and B look on at Speakers C and D, who are having a meeting to plan an event. They note that C and D have been working alone and that it is a lot of work. B says...]
  a. If I were them, I wouldn’t try to plan the whole thing by myself! [✓1 ant, 1 ana]
  b. If I were them, I wouldn’t try to plan the whole thing by themselves! [✓1 ant, 3.pl ana]

  (33) [Speakers A and B look on at Speaker C, who is holding a second place trophy. They note that C looks unhappy. B says...]
  a. If I were her, I’d be proud of myself! [✓1 ant, 1 ana]
  b. *If I were her, I’d be proud of herself! [✓1 ant, 3.sg ana]

  - The contrast between reflexive pronouns them and her is perhaps surprising
  - This seems to be a constraint on particular $\varphi$-features, and not reference

\(^{12}\)In other investigations to anaphors in shifted CID contexts (e.g., Lakoff 1996, Kamholz 2012, Kauf 2017), what is explored is the interpretation of anaphors that $\varphi$-match their closest syntactic antecedent. For example, they have explored sentences like ‘If I were you, I’d be looking at myself’, and whether it is a looking-at-addressee action or looking-at-speaker action. As far as I know, CID contexts with $\varphi$-mismatch on the bound anaphor have not been given serious investigation up to this point.
It is acceptable to have a mismatching antecedent for all uses of *them*
- Including plural-referring *them* in (32), as well as nonspecific *them* as in (34)

(34) If I were any of them, I'd be proud of *myself/them* selves

- Moreover, **the antecedent must be 1.SG** – any of the contexts above repeated with ‘we’ is unacceptable

(35) *Speaker A is going to the airport shortly, and asks Speakers B and C whether it’s a good idea to bring food or buy food on the plane. B replies on behalf of B and C...*

  a. If we were you, **we’d do our** selves a favor and bring food!  
     [✓ 1.PL ant, 1.PL ana]
  
b. *If we were you, we’d do **your** selves a favor and bring food!  
     [*1.PL ant, 2.SG ana]

(36)  

  a. If we were them, **we wouldn’t** try to plan the whole thing by **our** selves!  
     [✓ 1.PL ant, 1.PL ana]
  
b. *If we were them, we wouldn’t try to plan the whole thing by **them** selves!  
     [*1.PL ant, 3.PL ana]

- This data on the English CID contexts together yields three generalizations


<table>
<thead>
<tr>
<th>Some Preliminary Constraints on ϕ-Mismatch in English Counter-Identicals</th>
</tr>
</thead>
<tbody>
<tr>
<td>If (i) the antecedent of binding is a 1.SG pronoun, (ii) the mood is irrealis, and (iii) the reflexive self-phrase does not contain a 3.SG pronoun, then the pronoun and antecedent need not match ϕ-features to allow a grammatical reflexive interpretation in a CID context. If any of the conditions is not met, a mismatching reflexive pronoun is ill-formed.</td>
</tr>
</tbody>
</table>

- **There appear to be additional constraints on when mismatch is permitted between a reflexive pronoun and its antecedent in CID contexts (see Appendix B)**

### 3.5 Experimentally Confirming Constraints

- These native-speaker linguist judgments have been confirmed with a pilot study on Amazon Mechanical Turk task (*n=78*)
  - Each sentence was accompanied by a comic strip to provide context

---

Additionally, specific *them* (e.g., person whose known male/female gender identity is not being revealed), as well as gender non-binary *them* (i.e., person whose known gender is non-binary) seem to work fine here too, for speakers who accept these usages of *them*.
1. **SG antecedent with 2.SG reflexive pronoun:**

Should I buy food on the plane or bring my own?

If I were you, I'd do yourself a favor and bring food.

2. **1.PL antecedent with 2.SG reflexive pronoun:**

Do you guys think I should buy food on the plane or bring my own food?

If we were you, we'd do yourself a favor and bring food.

- Median scores so far are given below (1=“unnatural”; 5=“natural”)

(38) Ratings for φ-mismatch anaphors in conditionals

<table>
<thead>
<tr>
<th>Ant.</th>
<th>1.SG</th>
<th>2.SG</th>
<th>3.SG</th>
<th>1.PL</th>
<th>2.PL</th>
<th>3.PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG</td>
<td>–</td>
<td>4</td>
<td>2.5</td>
<td>–</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>1.PL</td>
<td>–</td>
<td>2</td>
<td>2</td>
<td>–</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

- There is a sharp change in the distribution of the ratings when the constraints on antecedents/pronouns are not met\(^{14}\) (Dark black lines indicate median score)

(39) Ratings of Mismatches, by Condition

- This corroborates the generalizations that 3SG pronouns cannot mismatch their antecedents, and that mismatch requires the antecedent to be 1SG

---

\(^{14}\)Further details of the data and analysis are given in the appendix. A full-fledged follow-up study is in progress.
4 Interim Summary

- **φ-features are visible at LF**
  - φ-compatibility with an antecedent is mediated by interpretation (cf. (26))
    
    (26) **Weak Identity Condition in English**
    The bound pronoun in an English reflexive must be able to be construed as (weakly) interpretive identical to its antecedent
    
    - e.g., *Each of us is behaving ourselves*

- **Anaphors are not a uniform class, morphosyntactically**
  - While some φ-bundles allow mismatch from the antecedent, 3.SG ones do not (cf. (37iii))
    
    (37) **Some Constraints on φ-Mismatch in English Counter-Identicals**
    If (i) the antecedent of binding is a 1.SG pronoun, (ii) the mood is irrealis, and (iii) the reflexive self-phrase does not contain a 3.SG pronoun, then the pronoun and antecedent need not match φ-features to allow a grammatical reflexive interpretation in a CID context. If any of the conditions is not met, a mismatching reflexive pronoun is ill-formed.
    
    - e.g., *We are behaving himself*

5 A Dual Approach to φ-Features

- Reflexive anaphors do not uniformly get their φ-features valued in the derivation
  - Premise: Reflexive anaphors structurally contain a (bound) pronoun
  - Premise: Not all (bound) pronouns are φ-deficient
  - Premise: Whenever a pronoun’s φ-features are derivationally valued, those φ-features must match the antecedent’s
  - Observation: Reflexive pronouns do not always match the φ-features of the antecedent
    
    ✷ ✷ Reflexive pronouns are not always φ-deficient

- Perhaps one could rescue a valuation-across-the-board approach by positing covert elements that give rise to the appearance of φ-mismatch, without φ-mismatch at a derivational level
  - i.e., something like the following:
    
    (35′) a. If I were them₁₈, I₅ wouldn’t try to plan the whole thing by my₅self!
    
    b. If I were them₁₈, I₅ wouldn’t try to OP₈ plan the whole thing by them₈selves!

- A covert-element approach would still require extra steps to explain the fact that 3.SG reflexive pronouns never mismatch their antecedents:

    (28) a. **Each of us** is behaving ourselves/%ourself/himself.
    
    b. **We** are each behaving ourselves/%ourself/%*himself.

    (33) a. If I were her, I’d be proud of myself!
    
    b. *If I were her, I’d be proud of herself!

    ◇ (Note: this is about π and #, and not about person alone; recall ‘them’)

    ✷ ✷ 3.SG reflexive pronouns require unique analysis
### 5.1 3rd Singular

- Notably, English 3.sg pronouns show the most $\varphi$-features of any person/number
  - The only pronoun that marks distinctions in gender, animacy, or genericity

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>me</td>
<td>us</td>
</tr>
<tr>
<td>2</td>
<td>you</td>
<td>them</td>
</tr>
<tr>
<td>3.M</td>
<td>him</td>
<td></td>
</tr>
<tr>
<td>3.F</td>
<td>her</td>
<td></td>
</tr>
<tr>
<td>3.INANIM</td>
<td>it</td>
<td></td>
</tr>
<tr>
<td>3.GENERIC</td>
<td>one</td>
<td></td>
</tr>
</tbody>
</table>

- Observation: English 3.sg pronouns are $\varphi$-valued for gender ($\gamma$), unlike other (pro)nominals

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ø</td>
<td>Ø</td>
</tr>
<tr>
<td>2</td>
<td>Ø</td>
<td>Ø</td>
</tr>
<tr>
<td>3</td>
<td>$[\pm F, \pm M]$</td>
<td>Ø</td>
</tr>
</tbody>
</table>

- Discourse-anaphoric 3.sg pronouns in English can, presumably, merge with specified $\gamma$-features, unconstrained by the derivation

- However, 3.sg *bound* pronouns\(^\text{15}\) behave differently
  - They do not allow $\varphi$-mismatch from their binder
    - (28) a. Each of us is behaving ourselves/%ourselves/himself.
    - b. We are each behaving ourselves/%ourselves/*himsel*.
  - (33) a. If I were her, I’d be proud of myself!
    - b. *If I were her, I’d be proud of herself!*

- This suggests their $\varphi$-features are *derivationally entangled* with those of the binder\(^\text{16}\)

**Idea:** 3.sg *bound* pronouns are subject to derivational constraints, due to $\varphi$-features

- 3.sg is unique in exponing $\gamma$-features
  - This suggests that $\gamma$-features trigger some process/constraint for $\varphi$-matching/valuation

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\gamma$</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>2</td>
<td>$\gamma$</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>3</td>
<td>$[\pm F, \pm M]$</td>
<td>$\gamma$</td>
</tr>
</tbody>
</table>

- In other person-number combinations, mismatch is possible: the $\gamma$-feature can remain unspecified, and the pronoun can go unconstrained by the morphosyntactic derivation

- This condition is purposefully stated in general, descriptive terms

---

\(^{15}\) I assume bound pronouns are identifiable by the grammar, due to a particular feature specification, e.g., the var feature discussed by Hicks (2009).

\(^{16}\) Perhaps because in 3.sg bound pronouns $\gamma$ is inherently unvalued (requiring syntactic valuation; cf. Sundaresan 2018’s 3rd-PERSON anaphor). Perhaps because $\gamma$ is valued (and there are constraints on where valued $\gamma$-features can appear).

\(^{17}\) Stated more formally: If $\gamma_{\text{pronoun}} \neq \emptyset$, then $\varphi_{\text{antecedent}} \subseteq \varphi_{\text{pronoun}}$.
5.2 Two Routes

- 3sg bound pronouns are unique: their $\gamma$-features must match the $\gamma$-feature of an antecedent
- In this way, there are two algorithms for determining whether a bound pronoun $\phi$-features are compatible with an antecedent
  - One route allows for $\phi$-mismatch between the pronoun and its binder (non-3sg)
  - The other disallows this, through additional constraints/operations on $\phi$-matching (3sg)

\[ (43) \]

\[ \begin{align*}
\text{build feature bundles} & \quad \text{in morphosyntax} \\
\gamma \neq \emptyset & \quad \text{no} \\
\text{Weak Identity Condition} & \quad \text{at LF} \\
\text{Special $\phi$-Matching} & \quad \text{Constraint(s)/Operation(s)} \\
\text{(only 3sg)} & \quad (\text{only 3sg}) \\
\emptyset & \quad (\text{LF? morphosyntax? vocabulary insertion?})
\end{align*} \]

\[ \text{\diamond (N.B.: this is not meant as a description of the formal derivation, but rather as a descriptive algorithm for the theoretician hypothesizing about derivations)} \]

- \(1\): Feature bundles are built during morphosyntax
  - e.g., \[
  \begin{bmatrix}
  \pi: 3 \\
  \#: pl \\
  \gamma: \emptyset \\
  \text{VAR: } x
  \end{bmatrix}
  \]
  , which spells-out as they/them/their and is interpreted as bound by $x$

- If, as is the case for 3sg bound pronouns, \(2\) $\gamma$ \textbf{is not} specified as $\emptyset$...
  - ...because, e.g., it has $\gamma$-features valued as $[-F, -M]$...
… then ⑴: invoke a separate set of constraints/operations

- Perhaps these are well-formedness constraints on possible φ-feature combinations between antecedents and bound pronouns (cf. (42))
  ◦ (In morphosyntax? At LF? Both?)
- It should be noted that prospects are weak for a syntax-only solution in which the antecedent nominal values the gender φ-feature in 3.sg pronouns
  ◦ Because φ-compatibility with an antecedent is mediated by interpretation (cf. the weak identity constraint and discussion in §3, including the cowgirl-type examples)

⑵: All bound pronouns are subject to the Condition on Weak Identity, (26)

- Necessarily at LF, since it is about construal and context

- Notably the constraints/operations in ⑴ don’t apply all the time
  - Allowing mismatch for bound pronouns that are 3.pl, 1st person, or 2nd person
  - Different bound pronouns are subject to different derivations, depending on the pronoun’s φ-feature specification in the morphosyntax

5.3 Deriving Well-/Ill-Formedness

- Let’s consider some data and state explicitly how the well-/ill-formedness for each is predicted

(44) The team will do it by itself

- Since y is specified for it, ⑴ is operative, and the φ-features of the antecedent match

(45)*/# They will do it by itself

- Since y is specified for it, ⑴ is operative, but the φ-features of the antecedent (namely pl) do not match, resulting in ungrammaticality and/or infelicity

(46) The team will do it by themselves

- Since y is not specified for them, all that matters is Weak Identity, which can be met in this context

(47)*/# He will do it by itself

- Since y is specified for it, ⑴ is operative, but the φ-features of the antecedent (namely m) do not match, resulting in ungrammaticality and/or infelicity

(48) # He will do it by themselves

- Since y is not specified for them, all that matters is Weak Identity, which cannot be met in this context: he cannot be construed as a group or as having non-binary/unknown/irrelevant gender, resulting in infelicity

(49) This person will do it by themselves

- Since y is not specified for them, all that matters is Weak Identity, which can be met in this context: this person can be construed as having non-binary/unknown/irrelevant gender

---

18 While changing pronoun reference for an individual is possible over the course of a conversation (cf. footnote 9), it appears not to be possible within a clause. This is probably akin to alternating between informal and formal pronouns for a single individual (an idea inspired by the work of Kirby Conrod). It is an open question as to what constrains there are about changing pronouns used to refer to an individual.
There are further constraints that have not yet been discussed

- e.g., When mismatch is possible/impossible between the pronoun and the √SELF nominal
  - (Recall data like ‘We all need to ask ourself a very serious question’)
  - Other constraints on mismatch in CID contexts are discussed in Appendix B

### 5.4 √SELF Number Mismatches

- So far: the pronoun can φ-mismatch its antecedent in a ways that have interpretive constraints/effects
- Briefly: the number features on the √SELF nominal can mismatch the number of the antecedent and/or the reflexive pronoun

\[
\text{(50) a. } \text{Each of us} \quad \text{has lifted the table} \quad \text{our} \quad \text{ourselves} \\
\quad [#:sg, \pi:3] \quad [#:pl, \pi:1] \quad [#:pl] \\
\quad b. \% \text{We} \quad \text{have all lifted the table} \quad \text{our} \quad \text{self} \\
\quad [#:pl, \pi:1] \quad [#:pl, \pi:1] \quad [#:sg]
\]

- Number mismatch between the antecedent/pronoun and the √SELF can yield particular interpretations

\[
\text{(51) We have all lifted the table ourselves/\%ourself.} \\
\quad \text{For those that allow both \textit{ourselves} and \textit{self} here, many people get a contrast in distributivity} \\
\quad \text{Namely that \textit{ourselves} likely gets a distributive interpretation, whereas \textit{ourself} likely gets a collective interpretation}
\]

- **Number on √SELF can have interpretive effects**

\[
\quad \text{More work is necessary to uncover the precise interpretive effects, and what constraints there are – such investigation is not taken up here}
\]

### 6 Conclusions

#### 6.1 Bound Pronouns and φ-(mis)match

- These constraints/generalizations we have seen are meant to be seen as descriptive explananda to be captured by a deeper analysis

\[
\text{(26) Weak Identity Condition in English} \\
\quad \text{The bound pronoun in an English reflexive must be able to be construed as (weakly) interpretive identical to its antecedent}
\]

\[
\text{(37) Some Constraints on φ-Mismatch in English Counter-Identicals} \\
\quad \text{If (i) the antecedent of binding is a 1.sg pronoun, (ii) the mood is irrealis, and (iii) the reflexive self-phrase does not contain a 3.sg pronoun, then the pronoun and antecedent need not match φ-features to allow a grammatical reflexive interpretation in a CID context. If any of the conditions is not met, a mismatching reflexive pronoun is ill-formed.}
\]

\[
\text{(42) Generalization on English Gender-Conditioned φ-Match} \\
\quad \text{If a bound pronoun has a gender φ-feature that is specified (i.e., not Ø), then the bound pronoun cannot have φ-features that conflict with those specified on the binder.}
\]
They lead to four larger conclusions

1. **Interpretation is involved in bound pronouns’ \( \phi \)-features**
   - Concepts like ‘be construed as’ in (26) are certainly interpretive
   - Social gender (continuously defined) matters, in ways different ways than grammatical gender (categorically defined) – even assuming gender manifests in English syntax
   - Contexts (such as CID contexts) can manipulate how \( \phi \)-features are interpreted

2. **\( \phi \)-features must be active at LF**
   - Assuming interpretation is involved in where mismatches are possible (and possibly involved in what \( \phi \)-features match with), this means these features are interpretable at LF
   - This contrasts with the view that is quite common amongst syntactic binding theorists:
     - “The form of the anaphor (e.g. the reflexive) plays no real role in the interpretation afforded […] This means that the agreement features are essentially bereft of semantic interpretation” (Drummond et al. 2011:399)
   - Instead, what has syntactic roots (\( \phi \)-features inside self-phrases) also has interpretational consequence (see also Heim 2008, Safir 2014)

3. **Morphosyntax builds reflexive anaphors**
   - \( \phi \)-feature bundles for bound pronouns are built in the morphosyntax
     - If English makes any use of bound pronouns that are \( \phi \)-deficient (e.g., \( \gamma \)-deficient) at the point of merge, such deficient pronouns are required only for 3.SG
   - How these \( \phi \)-features are valued influences the derivation and where mismatch is possible
     - e.g., Part (iii) of (37) has to do with the morphosyntax of \( \phi \)-features, as in (42)

4. **There are two ways for a bound pronoun to have well-formed \( \phi \)-features**
   - One does not require matching to the \( \phi \)-features of the nominal antecedent
     - (1st, 2nd, and 3.PL can behave this way)
     - These bound pronouns are candidates for merging with all \( \phi \)-features specified
       - (as must be possible for discourse-anaphoric pronouns)
   - The other requires more derivational machinery
     - (3.SG / \( \gamma \)-marked bound pronouns always behave this way)
     - Here, the derivation requires the pronoun’s \( \phi \)-features to match a binder; possibly because of \( \phi \)-deficiency that requires derivational valuation

---

**WE OUGHT TO REPLACE (2) IN ENGLISH WITH (26) AND (42)**

- Recall (2) suggests that English reflexive pronouns will *always* \( \phi \)-match their antecedent
  - This is true of only 3.SG reflexive pronouns
- Instead we need a weaker constraint
  - One about weak identity (26) applies to all reflexive pronouns
- 3.SG reflexive pronouns never violate (2) in English
  - They are subject to additional derivational constraints, loosely described in (42), because how its \( \phi \)-features (particularly \( \gamma \)) are valued
6.2 “Binding” Across Modules

* Properties of binding are not solely the product of syntax
  - Our investigation into $\phi$-features and nominal structures English reflexive pronouns has depended on multiple components of grammar
  - **Some syntactic**
    - Building English reflexive anaphors (pronoun + REFL + $\sqrt{\text{SELF}}$)
    - Building the $\phi$-bundles for the pronoun
  - **Some postsynactic** (based on syntactic input; see Ahn and Kalin *forthcoming*)
    - The case form of the reflexive pronoun
    - The (lack of) interpretation of $\sqrt{\text{SELF}}$ in reflexive self-phrases
  - **Some semantic/pragmatic** (based on syntactic input)
    - Determining whether a bound pronoun can be construed as weakly identical to an antecedent or not, on the basis of $\phi$-features
    - Semantic constrains mismatches (e.g., modality) in the CID contexts

* We need a multi-module approach to binding, whose name makes this obvious

  “Distributed Binding Theory”

* Other discussion of English-type reflexivity have also exposed that binding is done at multiple modules and at the interfaces
  - **Some of it is syntactic**
    - Reflexive features in the (extended) verbal projection (e.g., Labelle 2008, Kratzer 2009, Reuland 2011, Ahn 2015)
  - **Some of it is semantic/pragmatic** (based on syntactic input)
    - What types of meanings are possible for anaphors (e.g., Reuland and Winter 2009)
  - **Some of it is phonological** (based on syntactic input)
    - Where anaphors are prosodically weak/strong (e.g., Ahn 2015)

* This talk hasn’t aimed to produce definitive analyses for all of these problems
  - Rather, the goal is to show that binding does not emerge from a single grammatical module
  - And working on binding is interface work
References

Ackerman, Lauren. 2018. Syntactic and cognitive issues in investigating gendered coreference. ling-Buzz/004064.


Features, Identity, and ‘Yourself’


A  Rating Task Details

☆ Introduction to the task:

▶ “Here is a selfie that my astronaut friend took while doing a handstand on Mars next to the US flag”

▶ “Here is a selfie that my mom took a picture of herself while fishing”

☆ Target Stimuli:

▶ “If I were you, I’d do yourself a favor and bring food”

▶ “If we were you, we’d do yourself a favor and bring food”
“If I were you guys, I wouldn’t try to move the fridge by yourselves”

“If we were you guys, we wouldn’t try to move the fridge by yourselves”

“If I were her, I would be proud of herself”

“If we were her, we would be proud of herself”
“If I were you, I wouldn’t try to plan the whole thing by themselves!”

“I if we were you, we wouldn’t try to plan the whole thing by themselves!”

Summary of results:

T-tests confirm what can be seen visually above (data formatted as “T-value; p significance”)

<table>
<thead>
<tr>
<th></th>
<th>1pl-2pl</th>
<th>1pl-2sg</th>
<th>1pl-3pl</th>
<th>1pl-3sg</th>
<th>1sg-2pl</th>
<th>1sg-2sg</th>
<th>1sg-3sg</th>
</tr>
</thead>
<tbody>
<tr>
<td>1pl-2sg</td>
<td>-0.1; ns</td>
<td>-0.1; ns</td>
<td>-1.2; ns</td>
<td>-1.2; ns</td>
<td>-3.3; *</td>
<td>-3.1; *</td>
<td>-3.4; *</td>
</tr>
<tr>
<td>1pl-3pl</td>
<td>-0.1; ns</td>
<td>-1.1; ns</td>
<td>-1.1; ns</td>
<td>-1.1; ns</td>
<td>-3.3;</td>
<td>-3.1;</td>
<td>-3.4;</td>
</tr>
<tr>
<td>1pl-3sg</td>
<td>-1.2; ns</td>
<td>-1.1; ns</td>
<td>-1.1; ns</td>
<td>-1.1; ns</td>
<td>-3.3;</td>
<td>-3.1;</td>
<td>-3.4;</td>
</tr>
</tbody>
</table>

Pairwise comparisons using t tests with pooled SD (df=147)

- The 3 conditions deemed grammatical by native speaker linguists (1sg-2pl, 1sg-2sg, 1sg-3pl) were all given ratings significantly different from (in a positive direction) the other conditions
  - In addition, the ratings given to the 1sg-3sg condition was not significantly different from
the ratings given to the conditions with 1PL antecedents

- This corroborates the generalizations that 3sg pronouns cannot mismatch their antecedents, and that mismatch requires the antecedent to be 1sg

B Generalizations on φ-Mismatch in English Counter-Identifier

B.1 Review of the First Three Generalizations

1 The embedded clause must be in the irrealis mood

(52) a. If I₁ were you₈, I₅.as.₈ would bought myself₅.as.₈/myself₅.as.₈/yourself₅.as.₈ a new car.
   b. When I₁ was you₈ in a dream, I₅.as.₈ bought myself₅.as.₈/*yourself₅.as.₈ a new car.

- Perhaps the semantics of irrealis mood is necessary
- Perhaps it has to do with an operator in the syntax, introduced by irrealis mood

2 The antecedent has to be 1sg

(53) a. If I₁ were you₈, I’d do it myself₅.as.₈/yourself₅.as.₈.
   b. If we₁ were you₈, we’d do it ourselves₅.as.₈/*yourselves₅.as.₈.
   c. If you₁ were me₈, you’d do it yourself₅.as.₈/*myself₅.as.₈.

- Perhaps this has to do with semantic/syntactic privileges afforded to the speaker of the utterance

3 A 3sg reflexive pronoun can’t mismatch the antecedent

(54) a. If I₁ were her₈, I’d be proud of myself₅.as.₈/*herself₅.as.₈
   b. If I₁ were him₈, I’d be proud of myself₅.as.₈/*himself₅.as.₈
   c. If I₁ were them₈, I’d be proud of myself₅.as.₈/themselves₅.as.₈
   d. If I₁ were them₈, I’d be proud of myself₅.as.₈/themself₅.as.₈

- Recall that this is a constraint on syntactic φ-features and not number of referents: examples with themselves/themself are fine with singular referents

B.2 Three Additional Generalizations

4 The antecedent cannot be a local object(?), though passive subjects will do

(55) If I₁ were in your₈ shoes...
   a. I₅.as.₈’d ask the administrator to assign me₅.as.₈ to myself₅.as.₈/*yourself₅.as.₈
   b. I₅.as.₈’d ask the administrator to assign you₅.as.₈ to yourself₅.as.₈/*myself₅.as.₈
   c. I₅.as.₈’d ask the administrator to assign the task(s) to myself₅.as.₈/yourself₅.as.₈
   d. I₅.as.₈’d ask PRO₅.as.₈ to be assigned to myself₅.as.₈/yourself₅.as.₈

- Notice the interpretation here
  - In the irrealis clause, all pronouns/anaphors are interpreted in the same way: me-in-your-shoes
This yields maps onto a reflexive predicate: two identical (enough) arguments

Perhaps this is a condition on interveners?

That would mean that, in CID contexts, there are syntactically-specified features/operators that lead to the apparent mismatch

In this way, ‘me’ in (55a) intervenes between the reflexive pronoun and the feature/operator that allows it to have an interpretation of 5.as.8

Similarly for ‘you’, as in (55b), but 3rd person expressions do not intervene, as in (55c)

This sounds like 1/2 blocks long distance antecedents

Of course, the 1st person subject doesn’t intervene between the features/operators and the bound pronoun

So the relevant features/operators must be lower than subject position, in the middle-field, but not in the high left-periphery of the clause

This is a similar position to where modality might be specified; recall that irrealis mood is necessary for mismatch in CID contexts

Moreover, if me is interpreted as in the non-counterfactual sense, then mismatch is possible again

This suggests that the two pronouns in object position need to match only if both are interpreted counterfactually

Mismatch is impossible if there is another matching anaphor in the predicate/clause

This suggests some kind of grammatical constraint that, when the mechanism that allows mismatch applies, it applies uniformly

(Perhaps this is defined by the scope of some operator in the middle-field)

Adnominal ERs don’t allow mismatch, (58); even though agentive ERs and by-Xself do, (59)

This suggests some kind of grammatical constraint that, when the mechanism that allows mismatch applies, it applies uniformly

(Perhaps this is defined by the scope of some operator in the middle-field)

At the same time, it’s not that adnominal ERs must match the ‘real’ antecedent; it can match a local counter-identical pronoun

(60) If I were you...

a. I would describe you yourself

b. I would describe myself

But in those cases, the adnominal ER cannot mismatch its antecedent
(61) If I were you
  a. *I5.as.8 would describe me5.as.8 yourself5.as.8.
  b. *I5.as.8 would describe you5.as.8 myself5.as.8.

- Perhaps adnominal ERs (unlike adverbial ones) are too local to their antecedent, and no operators/heads intervene to yield a surface-mismatch

B.3 Suggestions Towards Understanding CID Constraints

- Counter-indexical contexts suggest that the mismatch is influenced by the syntax/semantics of mood and of scope/intervention
  - In particular, there appears to be some sort of operator in the inflectional middle-field or perhaps in the low left periphery in the verbal spell-out domain
    - This will cause everything to shift together
  - The inflectional middle-field seems a likely choice, given the facts about modality
    - Perhaps irrealis modals (that are in the middle-field) co-occur with the proper syntactic/semantic material to license apparent mismatch
- In this way, these CID contexts give the appearance of mismatch, but perhaps local match could be happening between the bound pronoun and the covert head/operator
  - Still unexplained: Why does this only happen with 1sg subjects? Why are the you and we subjects in (52) unacceptable?
- If there is local functional material that drives apparent mismatch in CID contexts, what blocks a 3.sg bound pronoun?
  - i.e., What about the covert material is incompatible with a 3.sg pronoun in the anaphor?
  - Perhaps it’s that the feature matching constraint described in (42) can’t be met with this silent material as a binder
    - (Perhaps this head/operator doesn’t have the γ-features necessary to match/value those of the 3.sg pronoun)