Features, Identity, and 'Yourself' **Byron Ahn** (Princeton University)

Introduction Reflexive nominals in English contain a pronominal component and a √SELF component:

(1) We will not embarrass our selves.

Both components of the anaphor and its antecedent in (1) match in the φ-features they expone (1.PL for *we* and *our*; just PL for *selves*), leading to analyses of binding that are predicated upon AGREE's feature-matching/feature-valuing (e.g., Heinat 2006; Reuland 2006; Rooryck and Vanden Wyngaerd 2011). Even in works that do not take AGREE to be at the heart of binding, it has been stated pervasively in the literature that, as a matter of linguistic fact, the antecedent and anaphor must match in φ-features (e.g., Hicks 2009; Kratzer 2009), and even has been formulated as a principle (Safir 2004's Antecedent Agreement).

In an overwhelming majority of cases, the following two generalizations hold: **(i)** Anaphors and antecedents match in features (as in (1)), and **(ii)** "The form of the anaphor (e.g. the reflexive) plays no real role in the interpretation afforded" (Drummond et al. 2011, 399). As such, building φ-matching into binding theory would seem unproblematic. However, **neither of these generalizations holds across the board**.

Internal Structure of English Anaphors Before exploring the above generalizations, we must establish that English *self* behaves as a nominal, and the pronoun that proceeds it is a separable possessor:

(2) We are self-deprecating. (≈ *"We deprecate ourselves"*) [Compounding/"Incorporation"]
(3) Mayor Embarrasses Self, Spouse (≈ *"The mayor embarrasses themselves and their spouse"*) [Headlinese]
(4) You should do it by your own damn lazy self. (≈ *"You should do it by yourself"*) [Adj. Modifiers]

These patterns are predicted with anaphors like *yourself* being built up in the morphosyntax, with the pronoun in the DP-layer and a √SELF morpheme in the NP-layer. Such analyses appear as early as Postal 1966, on the basis of data similar to (2). More recently, Ahn and Kalin (*forthcoming*) use data like (4) to argue the pronouns (e.g., *your* in *yourself*) are inalienable possessors in the syntax (and draw structural parallels to the wide variety of languages with anaphors comprised of possessors plus body-part nominals).

Core Data The data below show that φ-features of each of these three nominals (antecedent, pronominal, and *self*), while identical in most contexts, can also vary independently of one another:

(5) a. **Everyone** has been behaving **them selves** / %**them self**
3.SG 3.PL PL 3.PL SG
b. **Each of us** is doing it **our selves** / %**our self**
3.SG 1.PL PL 1.PL SG
c. **We** all need to ask **our self** [a very serious question] *–ABC Nightline, 1997*
1.PL 1.PL SG
d. **You guys** pushed **your self**, drove **your self**, sacrificed, trained, and competed
2.PL 2 SG 2 SG *–Mitt Romney, 2002*

(6) a. If I were you, **I** would do **your self** a favor and bring food. [*median rating: 4/5*]
1.SG 2 SG
b. If I were you guys, **I** wouldn't try to move the fridge by **your selves**. [*median rating: 5/5*]
1.SG 2 PL
c. If I were them, **I** wouldn't try to plan the whole thing by **them selves**. [*median rating: 4/5*]
1.SG 3.PL PL

The data in (5) come from corpus searches and native-speaker linguist judgments. (Data like (5a–b) are also discussed in Collins and Postal 2012.) The data in (6) —the literature (e.g., Lakoff 1996) appears to not yet have reported this sort of data— come from a sentence-rating task on Mechanical Turk (*n=78*). Each sentence was rated 39 times. (There were four different presentation lists; each list contained a different half of the sentences; each participant was randomly assigned to a list.) Each sentence was accompanied by a comic strip to provide context. Median scores are given to the right of each example (1="unnatural"; 5="natural").

Analysis Both pronominal and √SELF components of **English anaphors need *not* match their antecedent in φ-features**. As such, what rules out 'I_1 *like* it_1*self*' must not be a requirement on φ-feature matching. **Instead, binding requires *interpretive* co-identity.** (While '*they*' is 3.PL and '*each of us*' is 3.SG [as shown by verb agreement], they can be interpretively construed as a [bound variable] individual or as a group including the speaker, respectively.) Interpretational **co-identity between the bound pronoun and its antecedent is enforced at LF** (see, e.g., Kratzer 2009 for a technical implementation).

LF co-identity is straightforward in a (typical) case of matched φ-features; even in counterfactual contexts

like (6), feature matching is always possible ('...*I wouldn't...myself*'). Since co-identity is also possible when φ-features don't match, this LF identity function must not be particularly strict. This is coherent because φ-features (e.g., grammatical plural) don't align perfectly with interpretive concepts (e.g., notional plural).

It is *not* the case that the form of the anaphor "plays no real role in the interpretation"; in fact, further evidence (below) suggests LF plays a critical role.

The Role of LF With mismatches, only certain person-number combinations are possible. Compare (6c), with a 1.SG antecedent, and (7), with a 1.PL antecedent.

(7) *If we were them, **we** wouldn't try to plan the whole thing by **them** selves. [*median rating: 2/5*]

Whenever the antecedent was 1.PL and/or the anaphor's pronominal component was 3.SG, the sentences were judged unnatural in binding contexts. (The ratings for each of these unacceptable conditions was low [median rating ≤ 2.5], and pairwise t-tests reveal they are statistically different from the ratings of each of the acceptable conditions [i.e., those represented in (6)].) On the other hand, in contrast to the bound pronoun in (7), a non-bound 3.PL pronoun, as in (8), can occur in the modal clause with a 1.PL subject:

(8) If we were them, **we** wouldn't try to plan the whole thing without **their** boss.

(8) is only acceptable if '*we*' does not bind '*their*' (i.e., '*their boss*' is not interpreted as '*the boss of us-as-them*'). Thus it is feature (mis)match under binding that is constrained. The (un)grammaticality of feature mismatch in contexts like (6) or (7) is thus governed by the LF co-identity requirement for bound variable interpretations, and the interaction of counterfactuals and the interpretation of φ-features. (The precise reason for (7)'s unacceptability is not fully laid out here; what matters is the contrast between (7) and (8).)

Moreover, a "Freaky Friday" context (like (6) or (9)) with feature mismatch between the antecedent and pronominal component requires irrealis mood. Compare (6a) with (9):

(9) *When I was you (in my dream), **I** did **your** self a favor and brought food.

This further implicates LF (and not, e.g., AGREE) as what (dis)allows a φ-mismatched bound pronoun.

Finally, turning to √SELF, its number can independently contribute to interpretation. Consider (10):

(10) We all moved the piano by ourselves / %ourself.

Speakers report the '*ourself*' variant to strongly suggest a distributive reading ("*We all took turns and each moved it without help*"), while the '*ourselves*' variant is ambiguous. √SELF's φ-features also contribute to LF.

Conclusions The φ-features of antecedents, pronouns, and √SELFs can vary independently of one another. This is a natural state of affairs, if **what matters for binding is co-identity between the antecedent and pronominal component of the anaphor**, and this co-identity is enforced at LF (and not by AGREE).

On their own, the facts reported here provide a forceful empirical argument against the strongest form of analysis in which binding requires φ-feature matching. Beyond that, we can conclude that anaphors are syntactically complex, and the φ-features of the pronominal component, or of √SELF, don't just serve as markers of agreement; **manipulating these morphosyntactic features manipulates semantic interpretation**.

Finally, what has been viewed as a single phenomenon (matched φ-features between anaphors and antecedents) is necessarily multiple phenomena (co-identity between the antecedent and the bound pronoun; number on √SELF), which impact interpretation at LF. That is, **what has been thought of as a single operation/constraint (e.g., anaphor binding) ought to instead be broken down into multiple component operations/constraints**, and (at least) some of those components are deeply intertwined with LF.

Ahn, B., and L. Kalin. Forthcoming. What's in a (English) reflexive? Proceedings of NELS 48.

Collins, C., and P. Postal. 2012. *Imposters*. Cambridge, MA: MIT Press.

Drummond, A., D. Kush, and N. Hornstein. 2011. Minimalist construal: Two approaches to A and B. In *The Oxford handbook of linguistic Minimalism*. Oxford University Press.

Heinat, F. 2006. Probes, pronouns, and binding in the minimalist program. PhD Thesis, Lund University.

Hicks, G. 2009. *The derivation of anaphoric relations*. John Benjamins Publishing Company.

Kratzer, A. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *LI* 40.

Lakoff, G. 1996. Sorry, I'm not myself today. In *Spaces, worlds, and grammar*. Chicago University Press.

Postal, P. 1966. On so-called 'pronouns' in English. In *Report of the 17th annual round table meeting on linguistics and language studies*. Georgetown University Press.

Reuland, E. 2006. Agreeing to bind. In *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk*.

Rooryck, J., and G. Vanden Wyngaerd. 2011. *Dissolving binding theory*. Oxford: Oxford University Press.

Safir, K. 2004. *The syntax of anaphora*. Oxford: Oxford University Press.