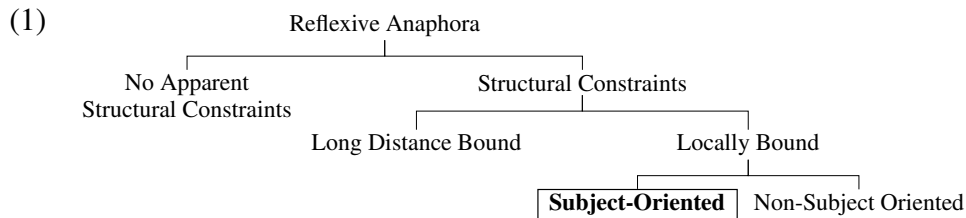


Reflexes of Reflexivity: Locality and the Interfaces

Byron Ahn
Boston University

This paper concerns itself with clause-bound reflexive anaphors which require the local subject as antecedent (Local Subject-Oriented Reflexivity; LSOR), boxed in the ontology of (1).¹



Many past approaches to reflexive binding are formulated in such a way that all reflexive anaphors are treated similarly. While it may be desirable in terms of formal simplicity for a single binding theory to govern all reflexive anaphors, this ought to be abandoned in favor of more complex solutions when necessary. In particular, LSOR anaphors are shown to be licensed through different syntactic derivations, indicating “reflexive anaphor” is not a homogeneous monolith in the grammar.

1 LSOR and Three Puzzles

Consider the unambiguous Shona (Bantu) sentence below:

- (2) Mufaro a- ka- zvi- bik -ir -a mbudzi [Shona]
Mufaro.1 SUBJ.1-PST-**LSOR**-cook-APPL-FV goat.9
'Mufaro_i cooked the goat_j for himself_{i/*j}.'

The *zvi-* prefix is a marker of LSOR, and so the only possible reading of the LSOR sentence in (2) is one where the beneficiary argument of *cook* in (2) is bound by the subject *Mufaro*, and not by the direct object *mbudzi*.² Clear morphosyntactic distinctions for LSOR clauses are found in a wide variety of languages and families, including Danish, French, Inuit, Japanese, Kannada, Lakhota, Russian, Shona, Sign Language of the Netherlands, and Tɔɔ sɔɔ; see Ahn 2014 for examples and citations. This cross-linguistically pervasive type of reflexivity leads us to our first puzzle, a naïve one:

(3) Naïve Subject-Orientation Puzzle

Why is it that special morphosyntactic marking of reflexives occurs only when the *subject* is the antecedent of the bound argument?

¹This ontology, inspired in part by Sportiche (2012), is intended to be descriptive, and it is almost certainly incomplete. There are different types of long-distance reflexives, and there are different types of exempt anaphors, possibly including (the different types of) logophors.

²Importantly, if (2) were embedded, the antecedent of the bound argument could *not* be the matrix subject.

In more formal terms, it is puzzling that morphosyntactic marking of reflexivity is sensitive to the grammatical role of the bound argument's antecedent, and that that grammatical role must be 'subject'. In more empirical terms, it is puzzling that languages do not exhibit special marking of reflexivity for when an object (as opposed to all other grammatical roles) is the antecedent of binding.

Though subject-orientation is a necessary condition on licensing LSOR, it is not a sufficient condition. Not *all* subjects can license LSOR; passive and raised subjects cannot license LSOR, as we will see. This raises a more empirically-informed (and more complex) puzzle.

(4) **Informed Subject-Orientation Puzzle**

Why can *only some subjects* license LSOR?

Additionally, there is a broad range of variation in how languages mark LSOR, leading to our final puzzle.

(5) **Typological Puzzle**

What is the range of possibilities for marking LSOR, and why?

The solution to all these puzzles is rooted in the syntactic derivation of LSOR, which is the consequence of two inter-dependent atoms of reflexivity: an anaphoric argument and a Reflexive grammatical voice.

(6) **The Core Underpinnings of LSOR**

- i. There is a unique grammatical Voice⁰ for reflexivity (REFL, which is associated with the semantic reflexivizer), and
- ii. An LSOR anaphor undergoes syntactic movement to VoiceP, whenever REFL (and hence the semantic reflexivizer) is merged.

The formal properties of these atoms (and of the Grammar) alongside variation in their lexical properties derives the solution to all of these puzzles.

2 LSOR and Theories of Binding

Well-established theories of reflexivity — both semantic ones (e.g., Bach & Par-tee 1980, Keenan 1988, Reinhart & Reuland 1993, Reuland 2011) and syntactic ones (e.g., Chomsky 1981, Hornstein 2001, Kayne 2002) — cannot (or do not) distinguish binding by a subject and binding by a non-subject from one another. In the frame of this paper, this might seem to be a short-coming; however, it has in fact been seen as a benefit: not all languages seem to differentiate LSOR anaphora from non-LSOR anaphora. For example, English employs a single set of anaphoric pronouns for LSOR and non-LSOR contexts:

- (7) a. Ken_k assigned Angie_j to herself_j.
b. Ken_k assigned Angie_j to himself_k.

This has led to a (perhaps implicit) line of reasoning in which reflexivity is an inherent aspect of language, but LSOR is not. As such, all languages share a base of rules governing the distribution of reflexives (which apply to LSOR and non-LSOR types of reflexivity), and LSOR —when it occurs— is the result of some additional, language-specific rule(s) and/or mechanism(s).

In particular, previous approaches have tried to derive LSOR from movement of the anaphor to near the subject position (e.g., Kayne 1975, Lebeaux 1983, Chomsky 1995, Safir 2004). Let us characterize the logic of such approaches as (8):

(8) **Reductionist Movement Approach to LSOR**

In an LSOR context, the anaphor undergoes movement from a thematic position to a position near to the subject (e.g. the INFL domain). The anaphor is necessarily bound by the subject, as a result of locality and normal binding conditions.

For example, Kayne 1975 argues that the subject-orientation of reflexive clitics in French is due to their movement to a preverbal position, where the only potential binder is the subject:

- (9) Jean₁ s'_{1/*2} est présenté les enfants₂ se. [French]
 Jean self PERF introduced the children self
 "Jean introduced the children to himself."

Safir (2004) also argues (based on the logic of Pica 1987) that Mainland Scandinavian anaphors, like Norwegian *seg selv*, are subject-oriented, due to covert movement of *seg* that targets a landing site that is local to the subject position. Chomsky 1995 (Ch.1) promotes a reductionist approach for binding more generally:³

"...the reflexive must move to a position sufficiently near its antecedent. This might happen in the syntax, as in the cliticization processes of the Romance languages. If not, then it must happen in the LF component."

Since reductionist movement analyses view conditions on LSOR to be a sub-case of general anaphoric binding conditions, this movement must be independent of (or additional to) the mechanism(s) for licensing reflexives.

A reductionist movement approach faces some challenges, regarding the timing of binding conditions. LSOR anaphors seemingly *must* get their reference *after* movement, otherwise movement wouldn't entail subject-orientation. If the a reductionist movement approach allowed binding of subject-oriented anaphors before movement, a subject-oriented anaphor like *se* in (9) would be able to have a non-subject antecedent bind it in the deep structure, counter to fact.

On the other hand, if all it takes to license LSOR is for the anaphor to move local to the subject, *any* subject should be able to license LSOR. Despite this prediction from a reductionist movement approach, derived subjects are well known to disallow LSOR (Kayne 1975, Burzio 1986, Rizzi 1986, Lidz 1996, Sportiche 2010, Storoshenko 2009), recalling our Informed Subject-Orientation Puzzle in (4). Kannada exemplify this in (10):

- (10) a. Hari tann-annu hoDe-du -koND -a [Kannada]
 Hari self -ACC hit -PST.PRT-LSOR-3SM
 'Hari hit himself.'
 b. hari (tann-age) santooshaagiruwaage kaNis-utt -aane
 Hari (self -DAT) be.happy seem -PRES-3SM
 'Hari seems (to himself) to be happy'

³Given that Romance has non-clitic anaphors that do not (appear to) move as well, it would seem to be that reductionist movement approaches such as the one promoted by Chomsky only apply to some anaphors.

- c. *hari (tann-age) santooshaagiruwaage kaNis-koLL -utt -aane
 Hari (self -DAT) be.happy seem -LSOR-PRES-3SM
Intended: 'Hari seems to himself to be happy'

When the (non-derived) subject *Hari* binds the anaphor *tann* in (10a), LSOR-marking with the verbal affix *-koL*⁴ is grammatical (and necessary; the anaphor *tann* requires it in this context). When *Hari* is the *derived* subject of the clause it in which it appears, as in (10b-c), the LSOR marker *-koL* is ungrammatical. (Note that the experiencer argument of *seem* in (10b) can be bound by a derived subject.)

Under a reductionist approach of movement, this would be surprising. All subjects reach subject position through A-movement (e.g. Koopman & Sportiche 1991), there is no clear reason for how 'normal' subjects and derived subjects (e.g. subjects in passive/raising clauses) would pattern differently with regard to allowing anaphor-movement to near the subject.

Instead, a reductionist movement approach to LSOR requires binding conditions to apply *both* before *and* after anaphor-movement. These constraints would have to be somehow relativized to LSOR, as non-LSOR anaphors are not constrained in timing in this way. They can be bound before all movements, (11a), or after all movements, (11b), or in intermediate positions, (11c).⁵

- (11) a. The crystal ball₂ predicted which stories about herself₁ Mary₁ would write ~~which stories about herself₁~~.
 b. The crystal ball₂ predicted which stories about itself₂ Mary₁ would write ~~stories about itself₂~~.
 c. Which stories about itself₂ did the crystal ball₂ predict ~~which stories about itself₂~~ Mary₁ would write ~~which pictures of itself₂~~?

(Note: the non-LSOR anaphors in (11) are not exempt anaphors.⁶) Additionally, languages which morphologically distinguish LSOR and non-LSOR anaphors (like French) support the generalization that lacking timing constraints is a property of *non-LSOR* anaphors (Charnavel & Sportiche 2013:(81,88), Starke 2001:(153)).

At this point, a reductionist movement analysis provides a tentative solution to our Naïve Subject-Orientation Puzzle and Informed Subject-Orientation Puzzle, but only with certain stipulations on the timing on binding for only LSOR anaphors. The descriptive fact that binding conditions for LSOR anaphors are timed differently from those for non-LSOR anaphors ought to be formally derived.

The basic idea of a reductionist approach – that LSOR anaphors move – finds support in the fact that islands affect on the distribution of LSOR anaphors.

⁴This morpheme has many allomorphs, such as the one in (10a), *-koND*.

⁵Or after reconstruction, if this is an independent operation that occurs after movement. If it is reconstruction, the puzzle changes to why LSOR anaphors must not reconstruct, whereas non-LSOR anaphors may. For a discussion in terms of reconstruction possibilities, see, e.g. Sportiche 2011.

⁶Exempt anaphors do not follow any binding conditions (e.g., Pollard & Sag 1992, Reinhart & Reuland 1993 and Reuland 2011). As such, the movement in (11b) is irrelevant, and **The crystal ball₂ predicted that someone₁ would write some stories about itself₂* would be (incorrectly) predicted grammatical. See Charnavel & Sportiche 2013 for how to distinguish exempt anaphors from other anaphors.

(12) **Descriptive Condition on Islands in LSOR Clauses**

In an LSOR clause, the bound argument must not be licensed in an island that excludes the subject.

French and Kannada islands (indicated by []s below) exemplify (12):⁷

- (13) a. Lucie s' est vu [French]
Lucie LSOR PERF seen
'Lucie saw herself.'
- b. *Lucie s' est compté(e) cinq [filles en dehors (de)]
Lucie LSOR PERF counted five girls outside (of)
Intended: 'Lucie counted five girls outside of herself.'
- c. Lucie a compté cinq [filles en dehors d' elle-même]
Lucie PERF counted five girls outside of herself
'Lucie counted five girls outside of herself.'
- (14) a. Hari tann-annu hoDe-du -koND -a [Kannada]
Hari self -ACC hit -PST.PRT-LSOR-3SM
'Hari hit himself.'
- b. *Hari [tann-annu mattu tann-a hendati-yannu]
Hari self -ACC and self -GEN wife -ACC
hoDe-du -koND -a
hit -PST.PRT-LSOR-3SM
Intended: 'Hari hit himself and his wife.'
- c. Hari [tann-annu mattu tann-a hendati-yannu] hoDe-d -a
Hari self -ACC and self -GEN wife -ACC hit -PST-3SM
'Hari hit himself and his wife.'

Even though the anaphor in (13c) and (14c) is bound by the local subject, the non-LSOR syntax (i.e. the non-LSOR anaphor *elle-même* in French⁸; no *-koL* suffix in Kannada) is required. It cannot be that LSOR anaphors like *se* simply have a property (like Kayne's (1975:Ch.5) +R feature) that requires its movement to a position near the subject because island effects arise in (14), where it is not the case that the anaphor used in LSOR clauses must always move to be near its antecedent (Lidz 2001, p.c.). (14c) is ungrammatical because the LSOR verbal suffix *-koL* is merged. At this point, there seems to be no reason to expect that islands would have an effect on whether a verbal affix can be present or not.

The solution to be pursued is that LSOR syntax (beyond the anaphor's featural properties) necessarily triggers the anaphor's movement to be near the subject.⁹

Movement analyses like (8) allow us to approach the naïve and informed subject-orientation puzzles, while deriving the island-sensitivity of LSOR anaphors. However, approaches like (8) must rely on stipulations on timing that apply only to

⁷The relevant argument occurs in an island in (13b&c): **Qui a Lucie compté cinq filles en dehors (de)?*. Similar data can be found with any number of islands (e.g. coordination, complex NP, etc.).

⁸Again, the issue of exempt anaphors arises here. See footnote 6.

⁹Lidz (2001) proposes a different analysis (what he calls Condition R) to account for this, which does not appeal to islands directly. Condition R may in fact be derived under the account to be presented here.

LSOR contexts, and do not provide a clear explanation for why islands block the merging of LSOR verbal suffixes. A complete theory must address these issues.

(15) **Desiderata for a Theory of LSOR**

- (i) to solve our Naïve and Informed Subject-Orientation Puzzles
- (ii) to understand why LSOR marking is impossible when the anaphoric argument is in an island separated from the subject
- (iii) to solve our subject-orientation puzzles without relying on timing-stipulations or needing to appeal to binding conditions

The key to meeting all of these desiderata is a correct understanding of what motivates the anaphor’s movement and the landing-site that this movement targets.

3 Reflexive Voice

This section describes a theoretical approach to LSOR which meets all the goals in (15). At its core, this theory invokes a secondary predicate, REFL, which is the head of VoiceP (which is in the extended verbal projection, outside of the Θ -Domain).

The REFL Voice and its semantic denotation of this REFL head are responsible for the locality and subject-oriented aspects of LSOR clauses, and license anaphors independent of any binding conditions necessary for non-LSOR anaphors. It does so by attracting an anaphor from within the Θ -Domain to its specifier, one which is featurally-specified in the appropriate way.

3.1 Subjects and Voice

Since the correct approach to LSOR needs to be able to refer to subjects, a basic question arises, which we have not yet addressed: which notion of subjecthood is relevant? The relevant notion could be the S-structure subject (S-subject) —the subject in grammatical subject positions, TP/IP— or the D-structure subject (D-subject) —the highest thematic argument. Consider the French data below (see e.g. Kayne 1975, Burzio 1986 and Sportiche 2010):

- (16) *Pierre_i se_i sera présenté par Jean_i [French]
 Pierre LSOR PASS.FUT introduced by Jean
Intended: ‘Pierre will be introduced by Jean_i to himself_i.’
- (17) *Tu_i te_i seras décrit par ta femme [French]
 You LSOR PASS.FUT described by your wife
Intended: ‘You_i will be described to yourself_i by your wife.’

Taken together, (16) and (17) demonstrate that, on their own, neither the notion of D-subject nor S-subject is enough. Instead, LSOR licensing depends on the subject both at S-structure and D-structure binding the reflexive argument. (Similar patterns are found in Kannada, Lidz 1996, and Shona, Storoshenko 2009.)

This fact can be summarized as the generalization below:

(18) **Generalization on Subjects in LSOR Licensing**

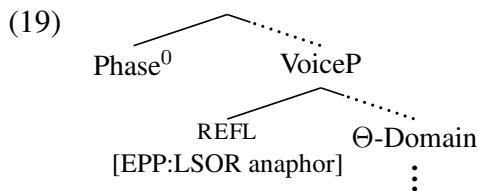
In cases of LSOR, the reflexive argument must be bound by the syntactic constituent which is both the D-subject *and* the S-subject.

This raises a question: how does the grammar enforce a restriction that requires that the D-structure subject gets mapped onto the S-structure subject? The answer lies in what controls mapping between D-structure and the S-structure subject: *grammatical voice*. We will assume a theory in which grammatical voice is instantiated syntactically as feature bundles merged as the head of VoiceP, which essentially acts as the gateway between the thematic domain and the subject domain (e.g. Ahn & Sailor 2014, Sailor & Ahn 2010).

As such, wherever Active Voice⁰ is merged, the D-subject becomes the S-subject, and any derivation in which the D-subject does not become the S-subject can only be generated through the merging of some non-active Voice⁰. For example, in verbal passives and raising over an experiencer clauses, the derived and demoted subjects reach their surface-structural positions as the result of a non-active Voice⁰ (see also Collins 2005a, 2005b).¹⁰

Since LSOR requires that the D-subject gets mapped onto the S-subject, and since the Voice⁰ determines which argument gets mapped on to the S-subject position, it follows that LSOR requires a specific Voice⁰ in the derivation: namely one which will not result in a derived subject. I will present evidence that this Voice is a Reflexive Voice⁰, abbreviated REFL. A derivation with REFL always yields LSOR properties. The fact that REFL is a Voice⁰ prevents LSOR anaphora in any contexts that requires a different Voice⁰, such as a passive clause, as REFL and PASS compete for the same position in the structure. (This is not to say that binding cannot happen in clauses without the REFL Voice⁰, such as passive clauses. This only predicts that the binding in such clauses cannot be of the *LSOR*-type.)

However, as Geniušienė (1987) aptly points out, “...the status of [reflexive predicates] with respect to voice is theory dependent in the sense that it depends on the definition of voice...” (emphasis mine). Thus in order to continue in trying to derive (18) with a Reflexive Voice⁰, let us define the Voice⁰ as a projection (just) outside of the Thematic Domain of the lexical predicate (e.g., Collins 2005b, Gehrke & Grillo 2009, Harley 2013), whose features ultimately determine what can become the grammatical subject (e.g., Sailor & Ahn 2010). The basic syntactic aspects of REFL are laid out visually in (19):



The syntactic feature that distinguishes REFL from other Voice⁰s is the EPP feature that attracts an LSOR anaphor. In terms of REFL’s semantic contribution, it is the source of the reflexive interpretation of the clause, denoting a function that coidentifies two arguments of the predicate:¹¹

¹⁰Empirical evidence from acquisition supports the idea that subject-to-subject raising over an experiencer predicates have a passive-like Voice⁰: verbal passives and raising over an experiencer are acquired rather late, and emerge simultaneously in child grammar (Orfitelli 2012).

¹¹It could be that REFL’s participation in the semantic derivation is responsible for “reflexive-marking” the predicate, in the sense of Reinhart & Reuland 1993 *et seqq.* See Ahn 2014.

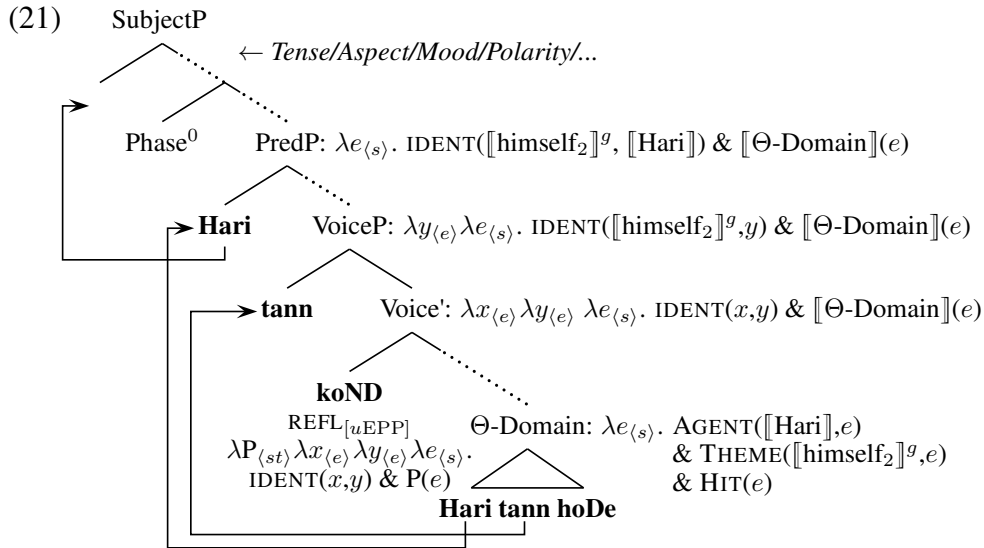
$$(20) \quad \llbracket \text{REFL} \rrbracket = \lambda P_{\langle st \rangle} \lambda x_{\langle e \rangle} \lambda y_{\langle e \rangle} \lambda e_{\langle s \rangle}. \text{IDENT}(x,y) \ \& \ P(e)$$

Note that no “valency-reduction” in the predicate, P , is necessary, nor is there any altering of already computed semantic values in P . Instead, the two arguments of IDENT , x and y , will always be the reflexive anaphor and the subject, for simple reasons of the mechanics of grammar.

The idea of a reflexive voice is generally supported by crosslinguistic facts. Briefly, a REFL Voice^0 is consistent with the fact that, when reflexives occur with special verbal morphology, that morphology tends to overlap with the morphology used in other grammatical voices (such as Passive, Medio-passive, Middle, Antipassive, etc.) in a many languages (e.g., Geniušienė 1987, Lidz 1996). Additionally, LSOR clauses may pattern with active clauses in some ways, and non-active clauses in others. This leads to the conclusion that LSOR clauses are not simply active voice clauses, nor are they passive or unaccusative clauses, either. Instead, LSOR clauses employ a separate, non-active, non-passive grammatical voice – Reflexive (REFL) Voice. We return to this in more detail in Section 4.

3.2 The LSOR Derivation

To more concretely present the effects of REFL as described in (19) and (20), the Kannada sentence in (14a) is derived below.



In (21), first *Hari* and *tann* are (external) merged in their respective thematic positions within the Θ -Domain of the predicate *hoDe* (which is dominated by VoiceP). Next, the anaphor *tann* is attracted to VoiceP by REFL 's EPP feature that probes for an LSOR anaphor, and it moves from its thematic position up to the specifier of VoiceP .¹² In this position, the anaphor composes with the IDENT function denoted by REFL .¹³ Following that, the D-subject *Hari* is attracted to the small-clause sub-

¹²This movement takes place in the narrow syntax. It is not LF-movement, as it has phonological effects. See Ahn 2014.

¹³The anaphor is interpreted as a simple pronoun (cf. Lees & Klima 1963 and Hornstein 2001). This makes the correct prediction that the morphological shape of the anaphor in LSOR could possi-

ject position (Bowers 2001's PredP), which is within the phase. (Harwood 2013 provides additional evidence for this Phase-internal position for subjects.) *Hari* is semantically local to the IDENT function of REFL Voice, and composes with it. (See Ahn 2014 for further motivation for, details of, and effects of this derivation.)

The technical details/decisions that comprise (21) need not hold for this analysis to be implemented. Instead, this REFL-based account of LSOR only relies on two, rather theory-neutral statements:

(6) The Core Underpinnings of LSOR

- i. There is a unique grammatical Voice⁰ for reflexivity (REFL, which is associated with the semantic reflexivizer), and
- ii. An LSOR anaphor undergoes syntactic movement to VoiceP, whenever REFL (and hence the semantic reflexivizer) is merged.

In this way, different frameworks and/or assumptions can be used to cover the same range of data, provided that they conform to these two statements. (See Ahn 2014 for some alternatives to the formalization in (21), which meet the description of (6), including one employing lambda abstraction, and one in which the LSOR anaphor is the reflexivizer.)

To be clear, the interpretation of the syntax in (21) in the semantic leads to a meaning where the subject in PredP is the binder of the LSOR anaphor in VoiceP, due to which arguments compose with IDENT. This allows us to meet all the desiderata laid out in (15), which we will very briefly discuss one by one.

For the first desideratum in (15i), the LSOR anaphor *be* interpreted as (relevantly) 'identical' to the (non-derived) subject, due to where each of them is (re-)merged in the syntax. Only subjects occur in PredP, and only non-derived subjects occur in PredP when Voice is REFL. (Derived subjects require a different Voice⁰, e.g., PASS.) For this reason, the IDENT function of REFL will only ever co-identify a non-derived subject and an LSOR anaphor. All other arguments in the thematic domain occur too low in the structure, and will not reach a point where it could compose with IDENT. Thus, binding between a direct object and an indirect object, for example, cannot employ REFL for semantic reasons.

Because this derivation relies on an LSOR anaphor moving to VoiceP, the second desideratum in (15ii) is entirely straightforward. REFL would never have its EPP feature checked, and IDENT would not provide any appropriate interpretation. Moreover, the reason verbal suffixes like *-koL* in (21) cannot merge when the anaphor is in an island, as in (14), is now understood. REFL induces movement of the anaphor, which is impossible when the anaphor is in an island. This is true even for derivations where the movement lacks effects on linearization and/or the moving anaphor is possibly phonologically covert.

Finally, this REFL-based derivation of LSOR properties does not rely on timing constraints or binding conditions, meeting the third desideratum in (15iii). Antecedent possibilities in LSOR are governed solely by (i) the positions of subjects, LSOR anaphors, and REFL Voice, and (ii) the syntax-semantics interface.

ble be the same as a pronoun – this is what is found in, for example, the 1st/2nd person of Romance languages, and even Old English (e.g. Siemund 2000).

4 Variation in LSOR Across Languages

4.1 Some Typological Patterns

Let us now consider what is *prima facie* a puzzle about the surface realization of LSOR across languages. LSOR can manifest in a given language with some number (possibly 0) of obvious morphosyntactic properties: unique anaphors, verbal affixes, and special word orders, in addition to other possible morphosyntactic effects. This has led to the puzzle presented in (5); to provide an adequate account of LSOR, *both* its formal properties *and* its language-specific realizations must be derived. We will now consider several generalizations that emerge after looking at a range of languages.¹⁴

Let us begin with a prediction of the REFL analysis. There are two atoms in this analysis —REFL and the LSOR anaphor it selects— and the former induces movement of the latter. None of these need have obvious effects on the surface string: REFL and the anaphor may (or may not) be silent, and syntactic movement (for various reasons) may (or may not) be string vacuous. This leads to 8 kinds of languages (without considering whether the LSOR anaphor has a unique morpho-lexical shape), only 6 of which are logically possible (if the anaphor is silent, it is not logically possible for movement to be visible).

Anaphor overt?	REFL overt?	Movement obvious?	Example language
Yes	Yes	Yes	Greek
Yes	Yes	No	Kannada
Yes	No	Yes	French, Czech
Yes	No	No	English, Tongan
No	Yes	No	Finnish
No	No	No	Shona
<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>(logically impossible)</i>
<i>No</i>	<i>No</i>	<i>Yes</i>	<i>(logically impossible)</i>

Table 1. Some Morpho-phonological Variation in LSOR at the Surface

All languages can be categorized in this manner.

In addition, since REFL only attracts certain anaphors, it must be that those anaphors have at least one probe-able feature that distinguishes them from other anaphors. This allows Vocabulary Insertion to identify LSOR anaphors as distinct from non-LSOR anaphors, on the basis of feature-bundles, predicting the possibility of a unique form for LSOR anaphors. At the same time, the lexicon need not supply a unique form for every distinct feature-bundle; there may be syncretism or homophony. Crosslinguistic patterns in Table 2 corroborate this analysis. (Note that non-LSOR anaphors aren't a homogeneous group, as Japanese shows; see (1).)

	French	Japanese	Czech	English	Tongan
LSOR anaphor	se	jibunjishin	se	themselves	kiautolu
Non-LSOR anaphor	eux-meme	{ jibun karejishin }	sebe	themselves	kiautolu
Non-Ref. Pronoun	eux	karera	je	them	kiautolu

Table 2. Variation in (3pl) Pronominals

¹⁴For reasons of space, primary data and their sources are not provided here; see Ahn 2014:Ch.5.

Similarly, languages are not uniform in the verbal voice morphology that occurs in LSOR clauses. Any individual Voice may be distinguished from or identical to other Voices, in terms of verbal morphology.¹⁵

	PASSIVE Voice⁰	MIDDLE Voice⁰	REFL. Voice⁰	ACTIVE Voice⁰
English	non-act. morph.	act. morph.		
Greek	non-act. morph.			act. morph.
Kannada	pass. morph.	refl. morph.		act. morph.
Finnish	N/A	mid. morph.	refl. morph.	act. morph.

Table 3. Morphological Realizations of Voice on the Verb

The pattern in Table 3 indicates that REFL Voice must be a unique category, to account for the cross-linguistic instantiations of REFL. Homophony and/or syncretism is responsible for the fact that LSOR markers may appear to be markers of a range of grammatical Voice-related functions (e.g. Geniušienė 1987, Lidz 1996).

Finally, REFL clauses may appear to more generally behave uniquely, as active-like, or as passive-like, in a range of other paradigms besides verbal Voice morphology. This is shown in Table 4.

LSOR clauses...	...pattern like actives	...pattern like non-actives	...pattern distinctly
Voice morphology	English	Greek	Kannada
Auxiliary selection	Spanish	French	?
TAM Markers	Mandinka	Kharia	?
Agreement Markers	Chickasaw	Lakhota	Shona

Table 4. LSOR effects on Morpho-Syntactic Paradigms

In each of the languages above, Voice has an established relationship with each of the relevant phenomena, indicating that Voice is what's operative in how LSOR influences these morphosyntactic domains. This reinforces the point made by Table 3; REFL instantiates its own category of Voice, and may have unique or shared paradigms with other Voices.

The Grammar does not require there to be any unique surface morpho-syntactic effects of LSOR – all surface effects of reflexivity could be silent, hidden, or homophonous with other lexemes (as seems to be the case in Tongan). What is critical is that the variation that does manifest across languages for LSOR is restricted to the surface forms of underlyingly distinct categories. In this way, the same abstract underpinnings is maintained across all Language in the structural representation of LSOR.

4.2 Typological Generalizations and Analysis

This data uncovers several generalizations, each of which shares (21) as a common source underlying all variation. The generalizations that have emerged are briefly summarized below.

¹⁵The way Table 3 is set up might implicate a kind of linear continuum of voices, with Passive and Active being diametrically opposed. This implication need not hold; e.g. Voice⁰s might be better described along multiple dimensions, and a linear representation based solely on “activity” is not adequate. (i.e. It is not clear how many features ought to be used to define Voice.)

First, no language has been found in which the anaphor has a special form predicated on the exact grammatical role of its antecedent, unless that role is the local subject. This is described in the following generalization:

(22) **Generalization on LSOR and Reflexive Anaphors**

If an anaphor requires its antecedent to have a certain grammatical role, then that grammatical role is that of the subject.

This descriptive empirical finding indicates that, while homophony may abound, the categories whose form the Grammar considers is constrained. The REFL derivation allows LSOR anaphors to be featurally distinct, but no feature bundle can define only local direct object oriented reflexivity to the exclusion of other types of reflexivity.

Second, though we find variation in the morphological realization of Voice, wherever we find a morpheme on the verb that is used for reflexivity, there is always a local subject binder (cf. Subbārāo 2012).

(23) **Generalization on LSOR and Reflexive Verbal Affixes**

If a verbal affix is used to mark reflexivity, the local subject must be the antecedent of binding.

There is no non-LSOR verbal morphology because non-LSOR is not constrained by a particular Voice, and non-LSOR binding can take place in any of them. That is, non-LSOR binding can occur in Passive or Active. However, LSOR is predicated upon REFL, so reflexive verbal morphology will only occur in LSOR contexts.

Lastly, REFL Voice may impact other morpho-syntactic phenomena, such as auxiliary selection, TAM markers, and agreement markers. Crucially, for all of the languages in Table 4, Voice has an established relationship with each of these phenomena.

(24) **Generalization on LSOR and Other Morphosyntactic Patterns**

If LSOR affects the realization of a morpheme in a paradigm other than the voice or anaphor paradigm (e.g. the aspectual paradigm), then voice more generally affects that morphological paradigm.

In other words, the relevant LSOR-sensitive morphosyntactic phenomenon is sensitive to grammatical voice more generally. This provides very strong evidence that reflexivity is formally represented in the same way as voice (i.e. as a Voice⁰).

All of these generalizations are the result of the formal derivation of LSOR demonstrated in (21), and the following theoretical generalization can be made:

(25) **Generalization on Possible Exponents of LSOR**

LSOR's morphosyntactic exponents are limited to Voice⁰ and its selectional relatives.

In other words, the only things that show morphosyntactic effects of REFL are Voice⁰, the specifier it selects (the LSOR anaphor), and other things that have more indirect selectional relationships with Voice⁰, such as agreement, aspect, and the auxiliary system (which are all structurally close to Voice). This reduces variation in LSOR marking to the selectional influences of Voice, which is itself constrained by locality.

Where there was once chaos there is order; all variation is in lexical items (cf. the Borer-Chomsky Conjecture, (Baker (2008))), and only on lexical items that can

be influenced by selectional properties of Voice. All syntactic properties of LSOR clauses remain constant across languages, because UG specifies (i) the height of REFL, and (ii) how its denotation necessitates movement. That means, languages as different as French and Shona exhibit the same LSOR effects, as we have seen. All languages encode LSOR, but finding its reflexes may require close investigation; even English exhibits unique grammatical properties in LSOR contexts, in its prosody; for reasons of space this will not be further addressed, see Ahn 2012,2014.

5 Conclusions

LSOR, the solution to the puzzles it implicates, all its properties, and apparent variation emerge from what UG provides: the general architecture of grammar, and the REFL Voice⁰. Recall (15), in which we laid out the desiderata for a complete theory of LSOR:

(15) **Desiderata for a Theory of LSOR**

- (i) to solve our Naïve and Informed Subject-Orientation Puzzles
- (ii) to understand why LSOR marking is impossible when the anaphoric argument is in an island separated from the subject
- (iii) to solve our subject-orientation puzzles without relying on timing-stipulations or needing to appeal to binding conditions

Each of these is addressed by the fact that (i) REFL, a unique grammatical Voice head associated with the semantic reflexivizer, is what licenses LSOR anaphors, (ii) REFL attracts an anaphor to VoiceP via selection, and that selection can impose restrictions on the kind of anaphor that is attracted, and (iii) due to the general nature of the organization of the grammar, subjects (and only certain subjects) and the LSOR anaphor are the only constituents capable of composing as co-arguments of an identity function introduced in the REFL Voice⁰. As such, subject-orientation is a core property of predicate-level reflexivization. It is not simply a special-case of normal binding conditions – defining LSOR in such a way would require unmotivated stipulations on the timing of binding, which in addition would ultimately be insufficient.

In addition to accounting for the desiderata in (15), this account predicts the range of morpho-syntactic variation in LSOR-marking, given entirely basic expectations on the kinds of lexical variation one ought to expect. Specifically, LSOR involves two lexical items (REFL and the moving anaphor), either or both of which may (or may not) have unique exponents, and REFL can share its morpho-syntactic paradigms with other Voice⁰s. It follows from this, that languages that do not obviously mark LSOR (e.g. English) ought to still employ REFL, despite lacking an obvious way of marking it – it's just that more careful investigation may be required to uncover its effects.

References

- Ahn, Byron. 2012. External argument focus and the syntax of reflexivity. *Coyote Papers: Working Papers in Linguistics, Linguistic Theory at the University of Arizona* 20.
- . *Giving Reflexivity a Voice: Twin Reflexives in English*. UCLA dissertation.
- , & Craig Sailor. 2014. The emerging middle class. In *Proceedings from the 46th Annual Meeting of the Chicago Linguistic Society*.
- Bach, Emmon, & Barbara H. Partee. 1980. Anaphora and semantic structure. In *Papers from the Parasession on Pronouns and Anaphora*, ed. by Jody Kreiman & Almerindo E. Ojeda, 1–28, Chicago Illinois. University of Chicago.
- Baker, Mark. 2008. The macroparameter in a microparametric world. In *The limits of Syntactic Variation*, ed. by T. Biberauer, 351–374. Amsterdam: John Benjamins.
- Bowers, John. 2001. Predication. In *The Handbook of Contemporary Semantic Theory*, ed. by Mark Baltin & Chris Collins, 299–333. Blackwell Publishing.
- Burzio, Luigi. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel.
- Charnavel, Isabelle, & Dominique Sportiche, 2013. Anaphor binding domains. ling-Buzz/001742.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- . 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Collins, Chris. 2005a. A smuggling approach to raising in English. *Linguistic Inquiry* 36.289–298.
- . 2005b. A smuggling approach to the passive in English. *Syntax* 8.81–120.
- Gehrke, Berit, & Nino Grillo. 2009. How to become passive. In *Explorations of Phase Theory: Features, Arguments, and Interpretation at the Interfaces*, ed. by Kleantes Grohmann. Mouton de Gruyter.
- Geniušienė, Emma. 1987. *The Typology of Reflexives*. New York: Mouton de Gruyter.
- Harley, Heidi. 2013. External arguments and the Mirror Principle: On the distinctness of Voice and v. *Lingua* 125.34–57.
- Harwood, Will, 2013. An existential crisis: understanding the aspectual restrictions on English existential constructions. lingBuzz/001829.
- Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Malden, MA: Blackwell Publishing.
- Kayne, Richard. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.
- . 2002. Pronouns and their antecedents. In *Derivation and Explanation in the Minimalist Program*, ed. by Samuel Epstein & Daniel Seely, 133–166. Malden, MA: Blackwell Publishers.
- Keenan, Edward L. 1988. On semantics and the binding theory. In *Explaining Language Universals*, ed. by John Hawkins, 105–144, Oxford. Blackwell Publishing.
- Koopman, Hilda, & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85.211–258.
- Lebeaux, David. 1983. A distributional difference between reciprocals and reflexives. *Linguistic Inquiry* 14.723–730.
- Lees, Robert, & Edward Klima. 1963. Rules for English pronominalization. *Language* 39.17–28.
- Lidz, Jeffrey. *Dimensions of Reflexivity*. University of Delaware dissertation.
- . 2001. Anti-antilocality. In *Long Distance Reflexives*, ed. by Peter Cole, Gabriella Hermon, & C.-T. James Huang, volume 33 of *Syntax and Semantics*. Academic Press.
- Orfitelli, Robyn. *Argument Intervention in the Acquisition of A-movement*. Los Angeles, CA: UCLA dissertation.

- Pica, Pierre. 1987. On the nature of the reflexivization cycle. In *Proceedings of North East Linguistic Society*, volume 17, 483–500.
- Pollard, Carl, & Ivan Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23.261–303.
- Reinhart, Tanya, & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24.657–720.
- Reuland, Eric. 2011. *Anaphora and Language Design*. Cambridge, MA: MIT Press.
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17.501–557.
- Safir, Kenneth. 2004. *The Syntax of Anaphora*. Oxford: Oxford University Press.
- Sailor, Craig, & Byron Ahn, 2010. The Voices in our heads: The VoiceP in English. Presented at Morphological Voice and its Grammatical Interfaces, University of Vienna.
- Siemund, Peter. 2000. *Intensifiers in English and German: A Comparison*. New York, NY: Routledge.
- Sportiche, Dominique, 2010. French reflexive *se*: Binding and Merge locality. ling-Buzz/001273.
- , 2011. The van Riemsdijk-Williams puzzle: in de Fourier’s footsteps. Presented at Parallel Domains: A workshop in honor of Jean-Roger Vergnaud. lingBuzz/001272.
- . 2012. Binding domains. Presented at Berkeley Linguistics Society 38th Annual Meeting.
- Starke, Michal. *Move Dissolves into Merge: A Theory of Locality*. University of Geneva dissertation.
- Storoshenko, Dennis. 2009. Investigating the Shona reflexive *zvi*. In *Selected Proceedings of the 39th Annual Conference on African Linguistics*, ed. by Akinloye Ojo & Lioba Moshi. Cascadilla Proceedings Project.
- Subbārāo, Kārumūri V. 2012. *South Asian Languages: A Syntactic Typology*. New York, NY: Cambridge University Press.