Anaphora and Soft Constraints*

Luigi Burzio, Johns Hopkins University

1. Introduction

In previous work, I have argued for the existence of several hierarchies controlling the behavior of anaphoric elements. It is easy to see, from the vantage point of "Optimality Theory," that hierarchies and soft constraints are just the same thing, yielding a direct translation of those earlier analyses into OT terms. In this paper I review the several hierarchies in question and argue that the OT analysis of anaphora they collectively entail has virtues of both empirical accuracy and cross-linguistic generality which are beyond the reach of any theory based on inviolable constraints.

2. Hierarchy 1: Referential Economy

In Burzio 1989, 1991 I have proposed that in a configuration in which there is coreference under C-command, standardly referred to as "binding," a reflexive element is preferable to a pronoun, which in turn is preferable to an R-expression. This hierarchy, given in (1) below, is interpretable as implementation of a principle of "referential economy," if one takes reflexives to have no inherent referential content, pronouns to have some, and R-expressions to have full referential content, as seems plausible.

(1) Referential economy: a » b » c

a. bound NP = reflexive
b. bound NP = pronoun
c. bound NP = R-expression

As is implicit in the fact that they are ranked, the constraints in (1) are violable, the examples in (2) illustrating some of the violations.

---

* The ideas presented in this paper were inspired in part by discussions with Colin Wilson, who presents related ideas in Wilson 1995. Much of this material was presented in a seminar at Johns Hopkins in the Spring of 1995. I am grateful to the participants in that seminar for numerous helpful suggestions and for stimulating discussion. Special thanks to Martha McGinnis for useful comments on an earlier version.

(2) a. John$_i$ saw himself$_i$.
b. John$_i$ expected Mary to see him$_i$.

The example in (2a) violates both (1b) and (1c), and the reason is that (1a), which it satisfies, is higher ranked. The one in (2b) violates the constraint in (1a), and the reason for that must lie in some other constraint of yet higher rank. We take the latter to be (3), where "locally" will await appropriate definition.

(3) A reflexive must be bound locally

The version of (2b) with a reflexive will now be taken to violate (3), while the one with the pronoun only violates the lower ranked (1a), thus representing the optimal candidate.

The system of (2a, b, c), along with (3), is the soft-constraint alternative to the principles A, B, C of LGB and later work. This system recalls earlier ideas by Bouchard, Pica, and Reinhart, as well as the "Avoid pronoun" principle of Chomsky 1981, but the central argument for it is novel. It is based on the two observations in (4).

(4) a. The two different morphological classes of reflexives, "SE" and "SELF" reflexives (in the terminology of Reinhart & Reuland 1993), differ in that "SE" but not "SELF" reflexives restrict the antecedent to subject and x-person.

b. Languages with SE reflexives (most of Romance, Germanic, Slavic) have pronouns which are locally bound by either objects or ¬x-person antecedents, as shown in (5).

(5) a. *Russian* (Timberlake 1979: 115)

Ja emu$_i$ skazal vse o nem$_i$ ...
I him told all about him
'I told him$_i$ everything about himself$_i$...'

b. *Italian*

Io$_i$ mi$_i$ vedo.
I me see
'I$_i$ see myself$_i$.'

c. *French* (Zribi-Hertz 1980: 137)

Victor$_i$ n'aime que lui$_i$.
Victor not loves but him
'Victor$_i$ only loves himself$_i$.'

The observation in (4a) will motivate the constraint in (6), where "x" varies cross-linguistically in ways to which we return.

(6) A SE reflexive must be subject and x-person bound.

The observation in (4b) will follow from taking (6) to outrank (1a), hence excluding reflexives in a subset of the local contexts, and consequently allowing the pronouns—the next best choice, in those contexts. The hard constraint alternative may superficially seem to handle these facts just as well, but that proves illusory. Let us take the traditional principles A and B as in (7).

(7) A. An anaphor must be α bound
B. A pronoun must be α free

In the most common type of formulation, motivated by English, α of (7a, b) is simply "locally." Evidently, when SE reflexives are at work, however, α must also include the restrictions of (6), as in (8).

(8) α = locally, subject and x-person

Given now that α appears not only in (7A), but also in (7B), it will follow that the corresponding pronouns will be allowed to be locally bound by either subjects or $\neg$x-person antecedents, correctly, as in (4b). It thus appears that there are two different means to express the complementary distribution of reflexives and pronouns described in (4). One is in terms of constraints whose domains of application may overlap, such as (1a), (3) (jointly referring to local binding), and (1b) (referring simply to binding), but which are ranked with respect to one another. The other is in terms of constraints which are unranked, but whose domains of application are complementary such as (7A, B). The solution in (7A, B) is incorrect, however, for the following reasons. What one observes, cross-linguistically, is that there is covariation of reflexive morphology and reflexive behavior, as stated in (6). This means that there is a cause-and-effect connection between those two facts, i.e. behind (6). In contrast, there is no covariation between pronoun morphology and their behavior. That is, unlike the reflexives, the pronouns of French, Italian or Russian are morphologically just like their English counterparts so far as anyone knows. Hence there cannot be any cause-and-effect relation there, which means that there is no reality to an (inviolable) principle like (7B), although the latter may serve as an accurate description. The noted complementarity of reflexives and pronouns must rather follow from the

fact that pronouns just fill the space from which reflexives are excluded, whatever that is, a fact only expressible by a soft constraint system such as (1) supplemented by (3) and (6).

Two points are worth considering here. One concerns known overlaps in the distribution of pronouns and reflexives. The system in (7) may seem better equipped to handle them, by having independent principles and hence greater expressive power. In fact, however, the proposed system is consistent not only with the rather general complementarity, but also with the specific overlaps that are found. As we see below, both the "locality" constraint (3) and the "subject and x-person constraint" (6) expand into constraint subhierarchies, each thus giving a family of "avoid reflexive effects." We will see that such families will have members ranked both above and below the "avoid pronoun" effect given by (1a). We then naturally expect that some members of the "avoid reflexive" families will rank comparably to the "avoid pronoun" of (1), yielding optimality ties between the pronominal and the reflexive options, whence the overlaps.

The second point concerns "long-distance" (LD) anaphora, and the prediction that our system makes that in languages that have long distance reflexives, pronouns should be long-distance free. That prediction is correct, as shown in (9a, b), both contrasting with their English counterparts.

(9) a. *Icelandic* (Anderson 1986: 73)

Jón$_i$ skipaði mér að raka	sig$_i$/*hann$_i$.
Jón ordered me that shave(inf)	self/him
'Jón$_i$ ordered me to shave him$_i$.'	

b. *Italian*

Gianni$_i$ videun	serpente vicino	a	sè/??lui$_i$.
Gianni saw a	snake near	to	SE/him
'Gianni$_i$ saw a snake near him$_i$.'			

This important fact has been almost entirely overlooked in the literature, in part because it is masked by the overlap of pronouns and reflexives in many LD contexts. As just noted, the latter follows as well from the present approach.

In sum, the distribution of pronouns is the "residue" created by the distribution of reflexives, a fact correctly captured by the soft constraints in (1), the hard constraints in (7) being mere descriptive artifacts.

3. Hierarchy 2: Optimal Agreement

The second hierarchy to consider has to do with the "x-person" restriction of (6). It appears true as a matter of cross-linguistic generalization that the instantiations of x as: first and second; third; and zero (=impersonal) stand in the implicational relations of (10).

(10) 1/2 → 3 → impersonal

That is to say, it is true that if a "SE" reflexive (in a language or a subsystem within a language) allows a first or second person antecedent, it will also allow a third person one, and in turn that if it allows a third person one it will also allow an impersonal one. An instance of the most restrictive case, allowing only impersonal antecedents is French *soi* (the non-clitic reflexive), whence the bound third person pronoun in (5c) above. The intermediate case, excluding first and second person antecedents, is represented by the reflexives of most of the Romance and the Germanic languages, whence the bound first person pronoun of (5b) above. The most permissive case, allowing antecedents of all persons, is that of Russian and other Slavic languages. In Burzio 1991, 1992, I argue that SE reflexives are always possible with impersonals because they have zero features just like impersonals, while their occurrence with non-impersonal antecedents is more restricted because it involves feature mismatches. The scale in (10), then, would indicate that third person is "closer" to zero features than first and second person. The hypothesis that impersonals and SE reflexives have the same features, which accounts for impersonals being on the end point of the implicational progression in (10) as just noted, also accounts for their frequent morphological identity, as for example with Romance *si/se*, both reflexive and impersonal. The Italian contrast in (11) is of further relevance.

(11) a. Gianni$_i$ pensa che [e]$_i$ vincerà.
Gianni$_i$ thinks that he$_i$ will win

b. Si$_i$ pensa che [e].$_i$ vincerà.
One thinks that he will win

In Italian, verbal inflection can function like a subject pronoun, leaving the subject position empty. In (11a), the embedded inflection is third singular, just like the main subject *Gianni*, whence the fact that the embedded subject can be understood as coreferential with it. The main inflection is also third singular, simply agreeing with *Gianni*. In (11b) on the other hand, the same coreference is not possible. This must mean that the main subject *si* is not third singular, or else the lower inflection, which is, should be able to corefer with it. But then the main inflection, which is itself third singular, must not be matching the subject's features. One conclusion is thus that the grammar must allow approximated agreement (between *si* and inflection, though not between *si* and a coreferential pro-

noun. See Burzio 1992). Since no other choice of inflection other than third singular is possible for the main verb in (11b) (the sentence being grammatical aside from coreference), the features of "impersonal *si*," which—as we just saw—are not third singular, must nonetheless be closer to third singular than to any other value of the inflection. The zero features hypothesis for "impersonals" seems the most natural in this respect, and is further supported by the fact just noted that impersonals are often also SE reflexives, once we observe that the latter are indeed featureless, witness their complete lack of morphological inflection for gender, person or number, unlike e.g. English reflexives. What follows from (11) is therefore: (i), that there must be approximated agreement (Burzio's 1991, 1992 "pseudo" agreement); and, in conjunction with other evidence: (ii), that both impersonals and SE reflexives must have zero features.

The implicational hierarchy in (10), true as a matter of observation as noted, is thus indeed interpretable in terms of approximated agreement, worse approximations always entailing the availability of better ones. That hierarchy translates directly into the constraint hierarchy in (12), which refers to antecedent-SE combinations.

(12) *1/2-SE » *3-SE » *impersonal-SE

In contrast, there is no translation of (10) (or any hierarchy), into a system of hard constraints.

4. Hierarchy 3: Blocks and Antecedents

As for the case of (10) above, it is also a matter of cross-linguistically observable fact, that different types of complements rank differently for their opacity to antecedent-anaphor relations, going from a maximum for indicative clauses, to a minimum for NPs. As discussed in Burzio 1996, the specific cause of the opacity appears to be not the clause itself, but rather the subject with which the relevant inflection is associated. The specific hierarchy of blocking elements in (13) then appears to hold.

(13) Subject of:

Indicative » Subjunctive » Infinitive » small clause » NP

Namely, it is true that if a language allows "long distance" (LD) anaphora into indicative complements (over their subject), as is the case in Faroese, it will also allow it out of subjunctives and all other complements on the right in (13), while if a language allows it out of subjunctives, it will not necessarily allow it out of indicatives, but it will allow it out of infinitives, small clauses and NPs, as for example in Icelandic. Similarly for the rest of the scale, which is instantiated by: Russian, allowing LD anaphora at most out infinitives; Danish, allowing it at

most out of small clauses; and Dutch, allowing it only out of NPs (see Burzio 1996 and refs. for finer-grained distinctions). In contrast to the subjects of (13), objects never block anaphoric relations, and can therefore be added as the rightmost point on the scale.

As argued in Burzio 1996, the scale in (13) pertains not only to blocking effects, but also to antecedency effects, in the sense that any case in (13) (e.g. subject of indicative) always turns out to be a better antecedent to a reflexive than any of the cases on its right (e.g. subject of infinitive). The parallelism between the notions of block and of antecedent is immediately visible in the subject/object asymmetry that enters into both: only subjects and not objects are possible blocks, just as subjects and not objects are possible antecedents to many reflexives. The parallelism is highlighted further by the fact that the exceptions to this basic generalization are the same in both cases: experiencer objects are subject-like both for blocking and for antecedency (see in particular Huang & Tang 1991). The differential behavior of antecedents is further observable in cases like (14) in Russian.[1]

(14) a. I on$_i$ ne prosil nikogo iz nix [provesti ego$_i$/sebja$_i$
and he not ask any of them lead him/SE

v nužnoe mesto...]
to needed place...

'And he$_i$ did not ask any of them to lead him$_i$ to the necessary place...' (Timberlake 1979: 124)

b. I on$_i$ stydilsja poprosit' kogo-libo iz nix
and he embarrassed ask any of them

[provestiego$_i$/?(?)sebja$_i$ v nužnoe mesto].
lead him/SE to needed place

'And he$_i$ was embarrassed to ask any of them to lead him$_i$ to the necessary place.' (Timberlake 1979: 124)

In general, the antecedent to a reflexive becomes critical when its prominence, measured on the scale in (13), approaches that of an intervening element as measured on the same scale. In (14a), the intended antecedent for SE is the subject of an indicative clause, while the potentially blocking element is the (less prominent) subject of an infinitive ('lead'). In (14b), however, the blocking element is still the subject of the same infinitive, but the antecedent is now itself the subject of an infinitive ('ask'), apparently not prominent enough. Putting aside

[1] Timberlake's (1979) diacritics "†, ?, *" are rendered here as "?, ?(?), *" respectively.

English-type reflexives for a moment, this suggests that the interpretation of reflexives is governed by the simple principle in (15).

(15) Reflexives must refer to the "optimal" antecedent

Optimal$_{def}$: most PROMINENT-AND-LOCAL

Regarding the fact that the more remote antecedent in (14b) and other similar cases is less than completely excluded, it seems plausible to suppose that "prominence" may also be contributed—to some moderate extent—by discourse (real or, for constructed sentences, imaginable). This would have the desired redemptive effect on the remote antecedent in (14b), and also account for the fact that, in cases like (14a), either antecedent, local or remote, appears possible. Each would be possible under different discourse conditions.

Another effect that (15) will correctly capture is the one described in (16). (For relevant examples and discussion, see Maling 1982, Giorgi 1984, Burzio 1996).

(16) LD reflexives can sometimes take object antecedents in local contexts, but never in non-local ones.

This effect simply follows from the fact that in a non-local context, an object (off the scale in (13)) will always be both less local and less prominent than the local subject, which will therefore be the optimal antecedent under (15). To the extent that it holds, the preference for subject antecedents will also follow from (15), given the greater prominence of subjects, assuming that an object is never relevantly more "local" than the local subject. As noted, experiencers appear to be more prominent than other objects, suggesting that the relevant notion of prominence is of a complex semantic nature, based in part on thematic structure, in part on the semantic content of the inflection the subject may be associated with, as expressed in (13), and in part on discourse semantics, as suggested just above.

LD anaphora is thus possible so long as a loss in the locality of the interpretive relation is offset by a gain in the prominence of the antecedent. This can be formalized by means of the constraints in (17), ranked as given.

(17) a. $NP_i^p \ldots SE_i \gg NP_i^{p-1} \ldots SE_i \gg \ldots \gg$

b. $* NP_i \ldots [_\alpha NP \ldots SE_i \ldots]$ (α=clause)

In (17), superscript *p* stands for the degree of prominence, as determined jointly by thematic role, discourse factors, and the semantic content of the associated

inflection in the manner of (13). The constraint hierarchy in (17a) thus imposes more prominent antecedents over less prominent ones, hence correctly excluding (under the relevant assumptions) object antecedents, as well as all remote antecedents which are less prominent than a local one. What remains to be excluded is remote antecedents which are only *as* prominent as a local one, and that is done by (17b). By taking the latter to rank below all members of (17a), LD anaphora will be possible as long as any "prominence" advantage exists, and correctly excluded otherwise.

Turning to the strictly local type of reflexives, it is well known that the distinction between them and LD ones correlates well with the morphological one between "SE" reflexives (LD) and "SELF" reflexives (local), as referred to in (4) above.[2] There are at least two candidates for what specific morphological property is responsible for their different behavior. One is the uninflected character of SE reflexives noted above versus the typically inflected character of SELF reflexives (MYself, YOURself,...), proposed in Burzio 1996. The other is the monomorphemic character of the former, versus the bimorphemic one of the latter, proposed by Pica (1991) and others. Whichever turns out to be correct, the effect of that property must be that of strengthening the locality effect of (17b), in a way expressible by a higher ranked counterpart like (18).

(18) *NP_i ... [$_\alpha$ NP ... $SELF_i$ ...] (α=clause)

If we take (18) to rank above all of (17a), the desired strict form of locality will result. Note that (17b) and (18) are just soft-constraint versions of Chomsky's 1973 "Specified Subject Condition," their ranks appropriately relativized both to reflexive morphology and to the morphology of the related inflection. We put aside here the "NIC" effect on reflexives, that is the fact that they (both SE and SELF types) are excluded as subjects of tensed clauses (see Burzio 1996). We will return to the lack of "subject orientation" of SELF reflexives below.

In sum, the interpretation of reflexives involves identifying an "optimal" antecedent, a task achievable by an appropriate system of soft con-

---

[2] Except for clitics, which are always both SE-reflexives morphologically, and strictly local. English permits some long distance uses of SELF-reflexives, as in (i), studied by Zribi-Hertz (1989), Reinhart & Reuland (1993) and others.

(i) It angered him that she tried to attract a man like himself

It is important to note that this does not represent a cross-linguistic generalization about SELF-reflexives. For instance, Italian *se-stesso* is completely impossible in (i). In contrast, I find the pronominal version *lui-stesso* marginally possible, suggesting the relevant structure in (i) may be *pronoun-himself* (parallel to *John-himself*), where the pronoun is null but recoverable from the features of the adjunct. This makes some sense of the LD relation, although the exact conditions of occurrence appear intricate and interesting. I will not be concerned with this case in the text.

straints—arguably, the one in (17a, b), (18). In contrast, conceptions based on hard constraints will fail in such a task, as they will necessarily fail to express the existence of hierarchical effects such as the one in (13).

5. Hierarchy 4: Avoid Structure

A fourth hierarchy relevant to the expression of coreference appears to rank at least four different types of elements in order of morphological complexity or "weight." The first type comprises elements like Italian *se-stesso*, French *lui-même*, English *his-own*, that is elements made up of a "pronominal" form and an intensifier adjunct. The "pronominal" here may be a true pronoun or a reflexive, the choice depending on the same principles ((1a, b), (3) above) that apply when there is no intensifier. The second type in the hierarchy is represented by the same pronominal elements, but without the intensifier. The next type is the class of clitics, which are morphophonologically "lighter" than full pronominals, and finally we find the "zero" case, that is no morphological expression of the coreferential element at all, be it as an empty category, or as no syntactic category—an issue which we will put aside. This hierarchy is given in (19) directly in the constraint version relevant for the discussion that follows.

(19) *full-INTENS » *full » *clitic » *ϕ

The ranked constraints in (19) can be taken to reflect a principle of economy of structure not unlike the one driving (1a, b, c) (and similar to one proposed in Cardinaletti & Starke 1994), but specifically referring to morphological structure. Since the more costly options in (19) are sometimes used, there must be some other principle at work to impose them. We take that principle to be some form of "recoverability." If the coreferential relation is fully recoverable simply from the semantic structure, i.e. if the semantics is "inherently" reflexive, then that relation need only receive minimal expression as imposed by (19). Conversely, if the coreferential relation is not recoverable from the inherent semantics, and—let us suppose—is in fact disfavored by it, then its morphological expression will have to be fully explicit. Among the contrasts that the above principle of morpho-semantic interaction (referred to as "Weak Anaphora" in Burzio 1994a) appears to account for are the Italian one in (20) and its English counterpart in (21).

(20) a. Gianni ϕ/*si apre gli occhi.
Gianni ϕ/to-self opens the eyes

b. Gianni si/*ϕ taglia i capelli.
Gianni to-self/ϕ cuts the hair

(21) a. Gianni opens his/*his own eyes.
b. Gianni cuts his/his own hair.

Here each of the (a) cases uses a form which can or must be lower on the morphological scale in (19) than the corresponding (b) case. The reason is that the cases in (a) are inherently reflexive, in the sense that one normally opens one's own eyes, opening someone else's having a completely different sense.[3] In contrast, "cutting one's hair" is not inherently reflexive since cutting someone else's seems to have roughly the same sense (e.g. one can do either with the same instruments). Now if the mapping from morphological categories into semantic conditions of occurrence was rigid, the two contrasts in (20) and (21) would remain completely unrelated to one another, since two languages map different morphological categories into the same semantic environments, requiring independent principles. It is only from the point of view of the scalar relations proposed that the two contrasts have something in common, which means that the scale exists. Soft constraints are then able to express that scale directly, as in (19).[4]

This apparatus also enables us to understand why English-type reflexives (SELF reflexives) are not subject oriented, unlike SE reflexives. Recall from the discussion in section 4 that object antecedents are disfavored, because semantically/ thematically less prominent. Such semantic bias will now invoke the hierarchy in (19), correctly requiring morphologically more complex elements for object antecedents. Note that it is clear that English reflexives qualify as morphologically complex relative to (19) since they are generally excluded (except for a few cases like (27g) below) in inherently reflexive contexts, as in **John shames HIMSELF* (i.e. 'is ashamed'), versus Dutch *Jan schaamt ZICH*, or **John mistakes HIMSELF* (i.e. 'is mistaken') versus German *Johann irrt SICH*. Conversely, they are possible in cases of semantically odd coreference in which SE reflexives are not, as in Jackendoff's 1992 "Mme. Tussaud's" cases, like *Ringo fell on HIMSELF* (i.e. 'on his statue') versus Italian **Ringo SI cadde addosso*. (See Burzio 1994a).

6. Interacting Hierarchies

The above discussion has argued for the existence of four different hierarchies at work. One of these, the one in (1), due to the principle of Referential Economy, will have an "avoid pronoun" effect (we put the Avoid R-expression effect

[3] In the "external" sense of the action, *his own* is in fact grammatical.

[4] It is not a problem for this discussion that inherent reflexivity cannot always be identified independently of the presence of a reflexive element. For instance, in the case of the structure *John mistakes self* in various languages, meaning "John is mistaken" one could claim (as has often been done) that the structure is simply intransitive (like its English translation). While semantically plausible, that view is excluded precisely because it cannot explain the reflexive. In contrast, an inherently reflexive analysis, no more implausible semantically (introspection provides no clue either way), explains the reflexive, as well as excluding non reflexive **John mistakes Mary*, etc., that exclusion being the only virtue of an intransitive analysis.

aside, here). On the other hand, two others will have "avoid reflexive" effects. The Optimal Agreement hierarchy (12) will have that effect when the antecedent-SE agreement is suboptimal. The Optimal Antecedent hierarchy based on (13) (and with objects added at the lowest end) will also have that effect since, when the intended antecedent is suboptimal, it will impose a different antecedent, in effect excluding the reflexive for the intended interpretation. The fourth hierarchy, related to the degree of inherent reflexivity (or irreflexivity) and given in (19) has an Avoid Structure effect. The three effects are summarized in (22).

(22) a. AVOID PRONOUN: Referential Economy (1)

b. AVOID REFLEXIVE: Optimal Agreement (12)
Optimal Antecedent (13)

c. AVOID STRUCTURE: *Morphological complexity (19)

It is clear that (22a, b) will be in conflict with one another, resolution then depending on the exact interdigitation of the two hierarchies into a single one. As for (22c), it will participate in that resolution as well whenever reflexive and pronominal choices differ on the scale in (19), as for example in (23).

(23) John had/took Mary with him/*himself

The structure in (23) is inherently reflexive in the sense that in 'x had/took z with y' an interpretation in which *x* and *y* are distinct is difficult or excluded. Such inherent reflexivity will require weak morphology, which in English only pronouns provide, whence the contrast in (23). In languages with appropriately weaker reflexives, e.g. Italian, the pronoun of (23) is sharply excluded, as imposed by Avoid Pronoun (22a). Hence the subhierarchy of (22c) also needs to be appropriately interdigitated with the rest, to yield violations of Avoid Pronoun such as (23) for the sake of satisfying Avoid Structure. The task of compiling an overall theory of anaphora from this perspective will thus essentially consist of appropriately integrating all individual subhierarchies into a single hierarchy. In contrast, and as I have repeatedly noted, the mere presence of hierarchies renders the hard constraint approach sharply inadequate. Note that existence of soft constraints is entailed even when the effects of the hierarchy are only observable cross-linguistically. Consider a hierarchy $C_1 \ldots C_i \ldots C_n \ldots$ like any of the ones noted above, taking the one in (13) for concreteness, which ranks complements for their opacity to (LD) anaphora. To say that such a hierarchy is universally true surely means that there is a sense in which it (or something that has that effect) exists in all languages. Now let us take a specific language, e.g. Russian, and consider that it only satisfies that hierarchy up to a C_i = subjunctive, infinitival complements and others being transparent to anaphora. What we will then have in Russian is a set of constraints: $C_{i+1} \ldots C_n$ which: (i) both exist (because universal) and (ii) are violated. The latter characteristic: both present and vio-

lated, is now just the definition of soft constraint (Prince & Smolensky 1993: 86). But in a number of cases that definition is in fact met directly within individual languages, without recourse to the cross-linguistic dimension. Thus, Referential Economy (1), if it exists as argued, is surely both present and violated in English given (2a, b). Similarly, the constraint enforcing subject orientation of LD anaphors is both present (in LD relations) and violated (in more local relations) in some languages, given (16) above. Also, a constraint excluding first and second person antecedents (of (12) above) appears to be both present and violated in Russian, given (24a, b, c).

(24) a. ... vy_i razrešite u vas_i/?(?)$sebja_i$ perenočevat'?
you allow by you/SE stay overnight
'... will you_i allow (us) to stay overnight with you_i?' (Timberlake 1979, 127, fn. 8)

b. $Starik_i$ ... prosil [na kurort $sebja_i$/ego_i pokuda ne
old man ask to resort SE/him now not

otpravljat'].
send off

'The $old\ man_i$ asked not to send him_i off to a health resort just now.' (Timberlake 1979, 124)

c. ... ja_i mog navredit' $sebe_i$/*mne_i ...
I_i can harm SE_i/ me_i
'... I can harm myself...' (Timberlake 1979, 113)

That person-constraint rears its head in (24a), where it is apparently combining with an independent Avoid Reflexive effect from non-optimal antecedency to overcome Avoid Pronoun, the two independent effects being otherwise only in a tie as in (24b). That same person-constraint is violated in (24c), however, where it is operating by itself. Similarly, Russian also violates the constraint of (13) against LD anaphora with infinitivals, as in (14b) above. Nonetheless the latter constraint shows its effects in (14a), where there is a weaker antecedent. In sum both presence and violation of a number of constraints is directly observable in individual languages.

We need to note that some of the above discussion raises questions that must await further study, however. One concerns that fact that in some cases optimization involves choosing the best form consistent with a specific interpretation, while in others it involves choosing amongst interpretations consistent with a fixed form. Thus, the Avoid Pronoun effect in (22a) comes from choosing the referentially most economical element consistent with a certain coreferential interpretation, while the Avoid Reflexive effect (22b) comes from choosing the

optimal interpretation for a form that has the reflexive. This means that constraints that need to interact with one another are in fact evaluating different sets of candidates, raising questions for the exact formalization of OT machinery. One important issue also concerns the "summation" effects such as the one of (24a), where both the Optimal Agreement and the Optimal Antecedent of (22b) combine to overcome Avoid Pronoun (22a). The question here is whether this entails abandoning or revising the "Strict domination" hypothesis of Prince & Smolensky 1993, that is the hypothesis that a higher ranked constraint will always prevail regardless of the number of violations of constraints of lower rank. In some cases this conclusion may be avoided in the face of "summation" effects, by supposing that a lower ranked constraint is only breaking a tie between higher ranked constraints rather than "teaming up" with another lower ranked one (essentially as proposed for (17) above, where (17b) was taken to determine the choice among antecedents of equal prominence, which would tie under (17a)).[5] Whether this solution is correct for (24a) (or even (17), for that matter) depends, however, on a finer-grained analysis than the present discussion can provide.[6] Note as well that, while the effects of agreement appear to combine with those of locality one way, disqualifying a more remote antecedent as in (24a), they do not appear to combine with it the other way, disqualifying a closer antecedent. That is, first and second person subjects appear to block LD anaphora just as much as third person ones, so far as I know.

Summing up again, a fully explicit theory of anaphora based on soft constraints will require resolving a number of issues of implementation. However, it seems unmistakably clear that soft constraints are at work given the noted hierarchical effects as well as the directly observable violations of constraints that otherwise need to be present in individual languages.

7. The Handle on Cross-linguistic Variation

On the proposed approach, a good deal of cross-linguistic variation will simply follow from differences in lexical inventories. For instance, we have seen that the bound pronoun of (23) above is made possible by the lack of SE reflexives in English. Similarly, the one of *John$_i$ read his$_i$ book* is due to the lack of reflexive possessives (in turn partially related to the lack of SE reflexives: Burzio 1996). The lack of SE reflexives also accounts for the contrast *The branch bends (*ITSELF)*, versus Dutch *De tak buigt ZICH*, Italian *Il ramo SI piega*, etc., as argued in Burzio 1994a. There are of course still many cases in which languages differ despite apparently identical inventories. We have seen above that languages with SE reflexives differ in the extent to which they permit LD anaphora. The distributional differences between Dutch *zich* and Icelandic *sig* are instructive, however. What we find in this regard is that Dutch *zich* has a consistently

[5] Thanks to Joel Hoffman for pointing this out.

[6] See also Burzio 1994c, which does not assume strict domination of constraints.

more limited distribution than Icelandic *sig* along several dimensions, and as indicated in (25).

(25) a. *zich* is more strictly local (discussion of (13) above).

b. *zich* is more strongly subject oriented (no cases of object antecedents as for Icelandic *sig*).

c. *zich* is more sharply excluded from local contexts which are not inherently reflexive, like 'John saw himself.'

What this suggests is a general "dispreference" for *zich*, not shared by Icelandic *sig*. This in turn suggests that the notion of lexical inventory is in fact itself not a rigid notion, some items being more central to the inventory than others. This conclusion seems natural, and consistent with our general perspective. Translating such "dispreference" for *zich* into a "*zich" constraint of appropriate rank then, and assuming further that summation of its effects with those of other constraints is possible, pending resolution of some of the issues noted earlier, the properties in (25) could follow unitarily.

Ultimately, some constraint re-ranking seems necessary, however. This is evident by comparing Icelandic *sig* and Russian *sebja*, whose respective distributions do not consistently represent subsets of one another. The Icelandic reflexive is more narrowly constrained relative to agreement (excluding first and second person antecedents), but less so relative to locality (permitting LD anaphora with subjunctive complements), showing that the relative ranking of locality and agreement constraints cannot be the same in both languages. While the exact extent of the possible cross-linguistic re-ranking must, like other issues, await further study, my own expectation is that it will turn out to be rather limited, much of the variation following from differences in inventory, both in the absolute sense (present/ absent in the inventory) and in the relative sense (present in the inventory but relatively dispreferred).[7] Independent syntactic differences, such as the different realization of possessor phrases in (20) and (21) above (dative object in Italian, versus possessive pronoun in English) will of course also end up playing a role.

In contrast, the patterns of cross-linguistic variation seem quite unmanageable for hard constraints, as we have already seen. The most direct form of the argument comes from the observation that any rigid mapping of morphological categories into domains of occurrence of the type once proposed in LGB lacks cross-linguistic generality. This can be seen by considering for example the specific domains employed in Reinhart & Reuland 1993 (RR), who attempt

[7] The latter sense corresponds in fact to a form of constraint re-ranking. E.g. a *SE constraint would be higher ranked in Dutch than in Icelandic.

a hard constraint analysis. One of these domains is defined by the notion "reflexive predicate," i.e. predicates whose arguments are coreferential (while not being inherently reflexive). Over this domain, RR's principles prescribe a SELF reflexive. While this is to some extent true, especially in English and Dutch, cross-linguistically all categories are in fact attested over this domain: SELF reflexives, SE reflexives, and pronouns, falsifying RR's claim:

(26) a. $John_i$ saw $himself_i$. (SELF)

b. *French* (Zribi-Hertz 1980: 137) (pronoun)

$Victor_i$ n'aime que lui_i.
Victor not loves but him
'$Victor_i$ only loves $himself_i$.'

c. *Italian*

Io_i mi_i vedo. (pronoun)
I me see
'I_i see $myself_i$.'

d. *Russian* (Timberlake 1979, 113)

... ja_i mog navredit' $sebe_i$/*mne_i ... (SE)
I_i can harm SE_i/me_i
'... I can harm myself...'

All of these cases follow from our approach, and were in fact discussed above: (2a), (5b, c), (24b).

Another of the domains targeted by RR's principles is that of inherent reflexivity, where either SE reflexives or no expression of reflexivity are required by their theory.[8] Here, again, all categories are in fact attested, including pronouns:

[8] RR's system of hard constraints in fact only prescribes reflexives over this domain, exclusion of SELF-reflexives being attributed to a principle of "economy" (fn.15). The latter is essentially a soft constraint, like "avoid structure" (19) above, witness its violation in (27g).

(27) a. *Italian*

Gianni apre gli occhi. (zero)
Gianni opens the eyes
'Gianni opens his eyes.'

b. John lost his cool. (pronoun)

c. John had Mary with him. (pronoun)

d. *Italian*

Gianni si vergogna. (SE)
Gianni SE shames
'Gianni is ashamed.'

e. *Frisian* (Everaert 1991)

Hy$_i$ skammet him$_i$. (pronoun)
He shames him
'He is ashamed.'

f. John is ashamed. (zero)

g. John behaved himself. (SELF)

Here, the cases in (a-c) were also discussed above and accounted for ((27b) is like (21a)), but our account is still a bit too crude to yield the full array. In particular, we have not explained why inherent reflexivity in the Romance languages gives a clitic for a direct object as in (27d), while it gives a zero for a dative possessor as in (7a). It also remains unclear why lack of SE reflexive results in zero in (f), SELF in (g), and a pronoun in (b, c, e). Additional factors (i.e. constraints) must therefore be postulated to handle the full set (though in some cases, e.g. (f) versus (g), the difference seems purely idiosyncratic). For a hard constraint approach, however, there is no reason for any of the variation at all: cases in which the prescribed elements are missing from the language should just be ungrammatical. Note too that the zero option of (f) arises only in languages that lack SE reflexives—another hierarchical effect, rather than a rigid principle.

One last domain relevant to RR's analysis is the complement of the previous two, essentially the domain of LD anaphora. By being subject to no constraints of its own, this domain is predicted to tolerate all three categories: SE reflexives; SELF reflexives, and pronouns. While the account of the two

previous domains was inadequate vis-à-vis the cross-linguistic facts by being too restrictive, the account of this last domain (their complement) appears inadequate by being too permissive. There are many cases within this domain where one or more of the three options are in fact excluded, as shown in (28). (See also fn. 2).

(28) a. *Italian* (??SELF/??pronoun)

Gianni$_i$ videun serpente vicino a sè$_i$/??se-stesso$_i$/??lui$_i$.
Gianni saw a snake near to SE/SELF/him
'Gianni$_i$ saw a snake near him$_i$.'

b. *Icelandic* (Anderson 1986: 73) (*pronoun)

Jón$_i$ skipaði mér að raka sig$_i$/*hann$_i$.
Jón ordered me that shave(inf) self/him
'Jón$_i$ ordered me to shave him$_i$.'

c. *Danish* (Jakubowicz & Olsen 1988: 6) (*pronoun)

John bad Peter$_i$ [PRO$_i$ anbringe [bøgerne
John asked Peter to put the books

bagved sig$_i$/*ham$_i$].
behind SE/him

'John asked Peter$_i$ to put the books behind him$_i$.'

d. *Dutch* (Everaert 1986: 235) (*SE)

Marie$_i$ maakte [mij jaloers op haar$_i$/*zich$_i$].
Marie made me jealous of her/SE
'Marie$_i$ made me jealous of her$_i$.'

e. *Dutch* (Everaert 1986: 212) (*SE)

Ria$_i$ vroeg ons [voor haar$_i$/*zich$_i$ te zorgen].
Ria asked us for her/SE to take care
'Ria$_i$ asked us to take care of her$_i$.'

From our perspective, the pronouns of (a, b, c) are excluded as violations of Avoid pronoun (22a), while the SE reflexives of (d, e) violate the combination of locality (17a, b) and the proposed **zich* specific to Dutch (or specifically

ranked for Dutch). The SELF reflexive of (a) violates (high-ranked) locality (18).

Grimshaw 1993 puts it well in stating that "Maximally general principles will inevitably conflict. The alternative is to formulate more specific principles that are designed never to conflict, and the price is generality." For a theory of anaphora that price is prohibitive. To achieve minimal empirical adequacy, hard constraints postulated for different languages would on the one hand have little to do with one another, and on the other would largely recapitulate independent facts about each language's lexical inventory.

8. Conclusion

In reviewing the distribution of anaphoric elements, one finds little cross-linguistic generalization in terms of which morphological class occurs where. The generalizations that do obtain are in terms of hierarchies such as the ones discussed in sections 1-4, which appear to be universal.[9] The fundamental implication of this is that an adequate theory of anaphora is bound to be one of ranked, violable constraints, of the kind proposed for phonology in Prince & Smolensky 1993, Burzio 1994c, although some aspects of its internal organization remain to be worked out. Similar conclusions appear to me justified for other areas of syntax, a view that other contributions to this volume will not fail to support. In Burzio 1994b, 1995 I argue that an account of cross-linguistic patterns of Case marking and NP movement also requires the use of soft constraints, while as early as Burzio 1986 (sect. 2.6.2) I employed constraint ranking to deal with difference in auxiliary selection in Italian and French in an analysis which still seems to me correct.

References

Anderson, Stephen (1986). The typology of anaphoric dependencies: Icelandic (and other) reflexives. In L. Hellan & K. Koch Christensen (eds.), *Topics in Scandinavian syntax*, Dordrecht: Reidel.

Bouchard, Denis (1982). The Avoid Pronoun principle and the Elsewhere principle. Ms., University of Ottawa.

Burzio, Luigi (1986). *Italian syntax: A government-binding approach*, Dordrecht: Reidel.

——— (1989). On the non-existence of disjoint reference principles. *Rivista di Grammatica Generativa* 14: 3–27.

——— (1991). The morphological basis of anaphora. *Journal of Linguistics* 27: 81–105.

——— (1992). On the morphology of reflexives and impersonals. In Christiane Lauefer & Terrell Morgan (eds.), *Theoretical analyses in Romance linguistics (LSRL XIX)*, Amsterdam: John Benjamins.

[9] The use of hierarchies in the theory of anaphora has been advocated at least as far back as Timberlake (1979).

——— (1994a). Weak anaphora. In G. Cinque, J. Koster, L. Rizzi & R. Zanuttini (eds.), *Paths towards Universal Grammar: Studies in honor of Richard S. Kayne*, Georgetown University Press.
——— (1994b). Case uniformity. Ms., Johns Hopkins.
——— (1994c). *Principles of English stress*, Cambridge University Press.
——— (1995). Anatomy of a generalization. To appear in the proceedings of the workshop on "Burzio's generalization." Utrecht, June 1994.
——— (1996). The role of the antecedent in anaphoric relations. In Robert Freidin (ed.), *Current issues in comparative grammar*, Dordrecht: Kluwer.
Cardinaletti, Anna, & Michal Starke (1994). The typology of structural deficiency: On the three grammatical classes. Ms., University of Venice and University of Geneva /MaxPlanck, Berlin.
Chomsky, Noam (1973). Conditions on transformations. In S. R. Anderson and P. Kiparsky (eds.), *A festschrift for Morris Halle*, New York: Holt, Rinehart and Winston.
——— (1981). *Lectures on government and binding*. Dordrecht: Foris.
Everaert, Martin (1986). *The syntax of reflexivization*. Dordrecht: Foris.
——— (1991). Contextual determination of the anaphor/pronominal distinction. In Jan Koster & Eric Reuland (eds.), *Long distance anaphora*, Cambridge University Press.
Giorgi, Alessandra (1984). Towards a theory of long distance anaphors: A GB approach. *The Linguistic Review* 3: 307–361.
Grimshaw, Jane (1993). Minimal projections, heads and optimality. Ms., Rutgers.
Huang, James & Jane Tang (1991). The local nature of the long-distance reflexives in Chinese. In Jan Koster & Eric Reuland (eds.), *Long distance anaphora*, Cambridge University Press.
Jackendoff, Ray (1992). Mme. Tussaud meets the binding theory. *Natural Language and Linguistic Theory* 10(1): 1–31.
Jakubowicz, Celia & Liz Olsen (1988). Reflexive anaphors and pronouns in Danish: Syntax and acquisition. Ms., Laboratoire de Psychologie Expérimentale, C.N.R.S., Paris and Linguistics Department, University of Copenhagen.
Maling, Joan (1982). Non-clause-bounded reflexives in Modern Icelandic. In T. Fretheim & L. Hellan (eds.), *Papers from the sixth Scandinavian conference of linguistics*.
Pica, Pierre (1984). Liage et contiguité. In J.C. Milner (ed.), *Recherches sur l'anaphore*, Paris VII.
——— (1991). On the interaction between antecedent-government and Binding: The Case of long-distance reflexivization. In Jan Koster & Eric Reuland (eds.), *Long distance anaphora*, Cambridge University Press.
Prince, Alan, & Paul Smolensky (1993). *Optimality theory: Constraint interaction in generative grammar*, Ms., Rutgers University and University of Colorado. To appear. Cambridge, MA: MIT Press.
Reinhart, Tanya (1986). Center and periphery in the grammar of anaphora. In B. Lust (ed.), *Studies in the acquisition of anaphora I*. Dordrecht: Reidel.
Reinhart, Tanya & Eric Reuland (1993). Reflexivity. *Linguistic Inquiry* 24(4): 657–720.
Timberlake, Alan (1979). Reflexivization and the cycle in Russian. *Linguistic Inquiry* 10(1): 109–141.

Wilson, Colin (1995) *Constraints on pronouns and reflexive predicates: Optimality theory and the syntax of reference*, Honors thesis, University of Colorado, Boulder, CO.

Zribi-Hertz, Anne (1980). Coréférences et pronoms réflechis: Notes sur le contraste *lui/lui-même* en français. *Linguisticae Investigationes* IV(1): 131–179.

——— (1989). Anaphor binding and narrative point of view: English reflexives in sentence and discourse. *Language* 65: 695–727.