

The Slovenian Orphan Accusative, Component Interfaces,
And Covert Grammatical Categories

Arnold M. Zwicky

Ohio State University and Stanford University

1. The Slovenian Orphan Accusative

Perlmutter and Orešnik (1973; hereafter P&O) observe that Slovenian exhibits all of the syntactic phenomena in (1) through (5), and they propose that these generalizations are nearly sufficient to explain the appearance in the language of a surprising construction they call the Orphan Accusative (OrphACC). The additional assumption needed to predict the OrphACC, in P&O's account, is rule ordering. In the remainder of this section I will illustrate the OrphACC and sketch P&O's analysis, which is couched in transformational terms. In the next section I observe that this analysis has several unfortunate properties, but that they vanish when the analysis is recast in nontransformational terms. However, the involvement of the grammatical feature of animacy in these phenomena turns out to be problematic. In sections 3-5 I shift from Slovenian to Russian and discuss the analytical and theoretical issues that arise there from the interactions of case, gender, number, and animacy.

- (1) The ACC form of the MASC SG is identical to the GEN form for +AN (animate) Ns, to the NOM form for -AN (inanimate) Ns; FEM SG Ns have distinct NOM, ACC, and GEN forms.
- (2) Modifiers - in particular, adjectives and determiners - agree with their head Ns in GEND, CASE, and NUM.
- (3) A definite pronoun can serve as an NP marking identity of sense as well as identity of reference.
- (4) A definite pronoun cannot serve as a modified N marking identity of sense, however; instead the N slot is empty when there are modifiers.
- (5) All definite pronouns, regardless of their reference, are grammatically +AN; in this respect they are like certain referentially inanimate nouns that are grammatically +AN, like *as* 'ace'.

Consider what happens when we construct a NP containing both an adjectival modifier and an identity-of-sense anaphor referring back to some earlier ACC SG N. According to (3) the anaphor can be a definite pronoun, and according to (5) such a pronoun will be +AN, but according to (4) it will not be realized phonologically. As for the modifier, what the remaining principles, (1) and (2), predict will depend on the GEND and AN values of the pronoun. If the pronoun is FEM, then (1) says it has a distinct ACC SG form, and (2) says that the modifier has the agreeing features CASE:ACC, GEND:FEM, and NUM:SG; these predictions are verified in (6c).

(6) -AN FEM ajda 'buckwheat':

- a. Katero ajdo hočete? 'Which buckwheat do you want?'
- b. Hočem navadno ajdo. 'I want ordinary buckwheat.'
- c. Hočem navadno. 'I want ordinary.'

If the pronoun is MASC, however, then (1) says that (since the pronoun is +AN, even for an inanimate referent) its form is GEN SG, and (2) says that the modifier has the agreeing features CASE:GEN, GEND:MASC, and NUM:SG; these predictions are verified in (7d), which contrasts with the ungrammatical (7c) - though (7c) is what we would expect from simply solving the analogical equation (6b) + (6c) = (7b) + X.

(7) -AN MASC ječmen 'barley':

- a. Kateri ječmen hočete? 'Which barley do you want?'
- b. Hočem navaden ječmen. 'I want ordinary barley.'
- c. *Hočem navaden. 'I want ordinary (=NOM).'
- d. Hočem navadnega. 'I want ordinary (=GEN).'

It is the form in (7d) that P&O identify as the OrphACC: a MASC (but not FEM) SG modifier in the ACC (but not any other) case which is 'orphaned' - that is, which is in combination with an empty N - and so has a special form, identical to the GEN. P&O's account of the OrphACC, which I have sketched informally above, depends not only on having the principles (1)-(5) in Slovenian, but also on several assumptions about the interactions among these principles, assumptions that were only implicit in my sketch. P&O, however, are quite explicit about these interactions. They assume three ordered transformations, which I paraphrase in (8): Pronominalization, corresponding to principle (3) but also incorporating a call on the lexicon, where P&O apparently assume principles (1), ACC Prediction, and (5), Animacy Prediction, apply; Agreement, corresponding to principle (2); and Pronoun Deletion, corresponding to principle (4). That is, pronouns are introduced as replacements for nominal constituents, and Agreement is determined with respect to these pronouns rather than the NPs they replace; having done their work with respect to Agreement, the pronouns are then deleted.

- (8) Pronominalization. A nominal constituent identical in sense to an antecedent constituent is replaced by a definite pronoun.
- Agreement. A modifier agrees with its sister nominal constituent.
- Pronoun Deletion. A definite pronoun is deleted when it is modified.

2. Component interfaces I

P&O's analysis predicts the OrphACC very nicely, but it has four aspects that are, to my mind at least, unsatisfactory. First, it seems to be intractably transformational; a nontransformational alternative is to be preferred if at all possible. Second, it posits a rule replacing anaphoric full NPs by pronouns, a step that is not easy to motivate even in transformational frameworks. Third, it relies on parochial (that is, language-particular) rule ordering; interactions predicted on universal principles are to be preferred wherever possible. Fourth, these parochial rule orderings include the stipulation that lexical insertion precedes Agreement: 'We are now proposing that the Orphan Accusative arises from the application of the rule of Concord at the stage of derivations at which the underlying head noun has been replaced by a pronoun.' (P&O: 427)

Ordering lexical insertion before the syntactic rule of Agreement is a particularly bad move, since lexical insertion (at least as P&O seem to understand it) makes available the full set of properties of lexical items: the values of features like AN, the choice of declensional paradigm, presumably even the constituent morphemes within the item and its phonological properties. That is, this part of the analysis makes it impossible to maintain sharp interfaces between the components of syntax, morphology, and phonology; but see Zwicky and Pullum (1986) and references therein for arguments that the autonomy of components should be maintained if at all possible. If the component boundary can be breached in this instance, then what sorts of interactions between syntax on the one hand and morphology and phonology on the other are excluded?

Fortunately, P&O's analysis of Slovenian can be translated into one that is free of the unsatisfactory aspects of the original - indeed, one that is fully consistent with the phrase structure framework of generalized phrase structure grammar (GPSG; see Gazdar et al. 1983). In such a framework there is no rule of Pronominalization; rather, pronouns are distributed freely in syntactic structures, subject only to local restrictions on their occurrence (and of course to a nonsyntactic requirement, that they must be semantically interpretable). Among the pronouns of Slovenian is an empty N, which I will assume has the features NI+PRO, +DEF, +NULL. This is no analytic innovation, since empty constituents of several types are now assumed in virtually all frameworks for syntactic description, including GPSG. There is then no Pronoun Deletion rule, but only principles distributing values of the feature NULL within branchings; one such principle disallows nominal constructs consisting of a [-NULL] modifier and a [+NULL] head.

The two aspects of their analysis that P&O treat as specifically lexical - ACC Prediction and Animacy Prediction - will be treated instead as syntactic principles, determining the values of CASE and AN, respectively, within a category on the basis of other features in that category (as Feature Co-occurrence Restrictions or Feature Specification Defaults, in the terminology of Gazdar et al. 1985). In particular, Animacy Prediction will require that an N with the features [+PRO, +DEF] also has the feature [+AN].

3. Covert grammatical categories I

The Slovenian analysis is still not trouble-free, however, since a family of problems surrounds the formulation of ACC Prediction. Thus far I have provided only informal characterizations of this principle, characterizations in which the FEM ACC, the MASC 'animate ACC' that is identical in form to the GEN, and the MASC 'inanimate ACC' that is identical in form to the NOM are systematically treated both as instances of a single grammatical category (ACC) and also as instances of three distinct grammatical categories (ACC, GEN, NOM). I will argue that the correct analysis does, in effect, have it both ways, but it is clear that in a nontransformational framework we cannot literally assume that an 'animate accusative' has both the feature CASE:ACC and the feature CASE:GEN in its syntactic description, for that would be contradictory. I have elsewhere (in Zwicky 1986b) argued that multiple feature marking should be countenanced in syntactic theory - but for the purpose of distinguishing inherent features from those imposed by rules of agreement or government, or of distinguishing impositions arising from different sources, and I cannot see that these proposals are applicable in the instance at hand.

I will begin, then, by considering analyses that choose one or the other of these feature assignments in the syntax. My discussion will use data from

(15)	+AN MASC 'old cat'	-AN MASC 'old table'
NOM	staryj kot	staryj stol
ACC	starogo kota	staryj stol
GEN	starogo kota	starogo stola

The analysis in (14) is a disaster from the theoretical point of view. Once again, the boundary between syntax and morphology would be breached. To get the right interaction between (14b) and (14c), with morphological realization preceding Agreement, either morphological realization must take place in the syntactic component, or Agreement must take place in the morphological component, or else the components as wholes must interact in exactly the opposite way from the one ordinarily assumed (in which syntactic rules are blind to the morphological composition of words, while morphological rules can be conditional on features distributed by syntactic rules).

Fortunately for component interfaces, (14) is simply wrong on factual grounds. There are clear instances of referral rules for Russian Ns, and in general these rules have no consequences whatsoever for the forms modifiers take. Thus FEM Ns ending in palatalized consonants have an ACC form that 'coincides with' the NOM (as Maltzoff (1984: 35) so carefully phrases it), but their modifiers nevertheless distinguish between ACC and NOM, as in the left column of (16). And MASC Ns ending in *g* have the declensional forms of the corresponding FEMs, including an ACC SG distinct from the NOM and GEN, but (as Klenin (1983: 9) observes) their modifiers nevertheless have syncretic realization, as in the right column of (16). It is also true that indeclinable Ns nevertheless have modifiers with full sets of declensional forms (as in the middle column of (16)), rather than an invariable form, as (15) would lead us to expect.

(16)	FEM 'old mother'	+AN MASC 'old attaché'	+AN MASC 'old uncle'
NOM	staraja mat'	staryj attaše	staryj djadja
ACC	staruju mat'	starogo attaše	starogo djadju
GEN	staroj materi	starogo attaše	starogo djadi

A variant of the analysis in (14) that requires no extraordinary component interfaces can be framed along the lines in (17). This approach allows a description of the facts in the first two columns of (16) - *mat'* can have the value NOM (when its GR is SU) or the value ACC (when its GR is DO) in the syntax, and *attaše* can have the full range of CASE values in the syntax - but it founders on the right column, since a DO *gdjg*- must receive the value ACC (so that its morphological realization can be distinct from the NOM and GEN) while its modifiers must receive the value GEN (because of their morphological realizations), thus contradicting the requirements of Agreement.

- (17) ACC Prediction is managed by syntactic rules distributing the value NOM, ACC, and GEN for CASE for BR:DO NPs.
- As in (11), except that some Ns (according to their values of GEN and AN) require the values NOM or GEN for CASE, rather than ACC.
 - As in (12).
 - Morphological forms are chosen on the basis of the values of CASE.

5. Covert grammatical categories II

I conclude that the correct account of ACC Prediction in Russian is more abstract than the ones in (14) and (17), which embody versions of the claim

that the CASE you see is the CASE you get. Consider instead the approach in (18), which uses a (more morphological) feature DECL distinct from a (more syntactic) feature CASE. For as like *staryj stol* are no problem in this sort of analysis. Both head and modifier are CASE:ACC and (because the N is GEND:MASC and +AN) DECL:2. The N *gdjg*- belongs to a morphologically exceptional subclass of Ns whose declensional forms are referred to the FEM, while the A *stac*- shows the default morphological forms for a word of DECL:2, including the referral of the ACC to the GEN. In a variant of this approach, outlined in (19), the feature AN is appealed to directly.

- (18) ACC Prediction is managed by syntactic rules distributing a (purely morphological) feature DECL of declension class.
- As in (11), with other syntactic rules determining the values DECL:1/2/3 on Ns according to their values of GEND and AN.
 - As in (12), except that modifiers also share the values of DECL on N.
 - Morphological forms are chosen on the basis of the values of DECL.
- (19) ACC Prediction is managed by syntactic rules distributing values of the (covert category) feature AN.
- As in (11).
 - As in (12), except that modifiers also share the values of AN on N.
 - Morphological forms are chosen on the basis of the values of AN.

From the theoretical point of view, both (18) and (19) are suspect, because they use the CAP and HFC to spread the covert categories DECL and AN, respectively - just the sort of use of noninflectional features that I spoke against in section 3. (Note that DECL, despite its name, is not inflectional in the technical sense; it conditions the choice of inflectional rules, but itself has no inflectional exponent.)

There are empirical problems as well, resulting from the fact that in these analyses genitive and animate accusative Ns do not constitute a natural syntactic class, but are related to one another only in the morphology. As it happens, however, there is at least one place in Russian syntax where [CASE:ACC, GEND:MASC, +AN] groups with [CASE:GEN] and the other oblique cases (DAT, PREP, INSTR), as against the direct cases [CASE:NOM], [CASE:ACC, GEND:NEUT], [CASE:ACC, GEND:FEM], and [CASE:ACC, GEND:MASC, -AN]. The cardinal number words 'two' through 'four' govern CASE:GEN and NUM:SG within NPs in direct cases, but within NPs in oblique cases they agree in CASE and NUM:PL with their heads (see Zwicky (1985b: sec 6.3) for a GPSG treatment of these and related facts). As a result, 'three cats' looks throughly PL (as well as genitive) in the ACC, while 'three tables' has a clearly SG head in the ACC, as in (20). But to state the generalization about CASE and NUM government with cardinal number words, we need to treat the syncretic ACCs that look like GENs as forming a class with the true GENs, which is not possible with the assignment of features used in (18) or (19).

(20)	NOM	tri kota	'three cats'	tri stola	'three tables'
	ACC	tr ^g x kotov		tri stola	
	GEN	tr ^g x kotov		tr ^g x stolov	

Clearly we need to have it both ways. In some ways MASC SG ACCs are distinct from GENs and NOMs, but in other ways the +AN ones are the same as GENs (and the -AN ones the same as NOMs). I propose to treat these cross-cutting assignments of forms to classes in the syntax as exactly parallel

to cross-cutting assignments of segments to classes in phonology. That is, I will decompose the values of the feature CASE into sets of features, thus splitting ACC into several subCASEs.

As a formal move, this has all the advantages of the analysis in (18) using the feature DECL, but does not involve spreading a covert category and permits the direct/oblique distinction to be made fairly simply (as in (23) below). The proposal is outlined in (21), and the roles played by the new features, X and Y, are specified by the rules in (22); note that (22c) says that the value of Y is closely related to, but not identical to, the value of AN. The rules in (22), together with the morphological defaults in (21c), correctly describe all of the facts about ACC Prediction listed earlier in (13).

(21) ACC Prediction is managed by syntactic rules determining the values of X and Y in CASE of N[CASE:(ACC)] according to the N's values of GEND and AN; see (22).

- a. The default value of CASE for NP[GR:DD] is (ACC); there are three subCASEs, {ACC, +X, -Y}, {ACC, -X, +Y}, and {ACC, -X, -Y}.
- b. As in (12).
- c. Morphologically, the defaults are for the first of these subCASEs to be realized via the distinctly ACC forms, the second by referral to GEN, and the third by referral to NOM.

- (22) a. If N is NUM:SG, GEND:NEUT, CASE:(ACC), then it is CASE:(-X, -Y).
- b. If N is NUM:SG, GEND:FEM, CASE:(ACC), then it is CASE:(+X, -Y).
- c. The default is for @AN, CASE:(ACC) N to be CASE:(-X, @Y).

(23) The direct CASEs are NOM and (ACC, -Y); all others are oblique.

The decomposition of CASEs into features, which plays such an important role in my analysis, is no cheap formal trick. Such a decomposition is called for in a large number of other instances. It is, I believe, the appropriate mechanism for stating that in Russian the prepositions alluded to in (9b) govern either ACC or PREP, depending on their meaning; syntactically, ACC and PREP should share a feature (call it +SPAT), so that the rule in question stipulates that objects of these prepositions are +SPAT, the objects of other prepositions being specified CASE:(+SPAT, +ACC) or CASE:(+SPAT, -ACC) or some other CASE entirely. Presumably, decomposition of CASE is also an appropriate method for dividing the CASEs of Russian into a direct and an oblique subset, +OBL being the default value of the feature in question.

Feature decomposition of CASE is also the natural way to describe the marginal or sporadic CASEs of many languages, for instance, PART and LOC in Russian and what I will call IGEN in English. Russian PART is a special set of forms used with partitive meaning, and it is available only for certain MASC nouns; otherwise GEN is used for partitives (Mal'tzoff 1984: 28f). Russian LOC is a special set of forms used with locational meaning, and it is available only for certain MASC nouns serving as objects of the two prepositions y and na; otherwise PREP is used for locationals (Mal'tzoff 1984: 30f). English IGEN is a special set of forms used with predicate possessives and possessive objects of the preposition of (IDis book is @IGEN, a book of mine), and it is available only for the personal pronouns; otherwise GEN is used for possessives. In each such instance, we can say that the marginal CASE shares one feature with its default counterpart but differs from it on another

feature: CASE:PART = CASE:(+GEN, +PART), CASE:GEN = CASE:(+GEN, -PART), for instance. Then if rules for the default CASE are stated in terms of the shared feature they will cover the marginal CASE as well, unless there is a stipulation specifically to the contrary.

6. Concluding remarks

To sum up: My proposal treats what are sometimes, rather awkwardly, called the 'animate accusative' and 'inanimate accusative' of Russian, Slovenian, and other Slavic languages (as opposed to the plain 'accusative' exhibited by FEM SG Ns) as subCASEs of ACC, a move with parallels elsewhere in Russian and in many other languages. The analysis outlined in (21)-(23) then describes the facts of Russian without violating strong universal hypotheses about the interfacing of grammatical components and about the role of covert grammatical categories in syntactic rules.

One lesson to be drawn from this discussion is that we must insist as much as possible on having precise statements of grammatical rules, located within an explicit framework of assumptions. Truly formidable analytic problems, as well as central issues of theory, may lie concealed within informal statements like the Slovenian ACC Prediction rule in (1) or its more detailed Russian counterpart in (13). And traditional scholarship may give little hint of these complexities: 'It is a curious fact that questions of grammatical agreement which often baffle the non-native speaker tend to be treated in an offhand manner in Russian grammars and have not attracted much scholarly attention to date.' (Crockett 1976: 1)

Another lesson is that it is easy to underestimate the extent of grammaticization in particular languages, and indeed in language. The first analyses I considered for Russian were attractive largely because they embodied the principle that the CASE you see is the CASE you get, a principle that directly reflects the central sound-meaning function of systems of agreement, according to which phonological identity signals grammatical relationship. It might be that systems of agreement arise, both diachronically and ontogenetically, to serve this function directly. But it seems that they become grammaticized, indeed syntactified, with lightning speed. Despite occasional appearances to the contrary, agreement systems do not seem to involve phonological or morphological copying, but instead are universally matters of syntactic feature sharing - a position that is in fact assumed without argument in the thoughtful crosslinguistic survey of agreement phenomena by Moravcsik (1978).

The evidence from Slavic suggests that fairly complex systems of grammatical agreement can be remarkably stable, once established through the side-effects of phonological change, through language contact, or whatever. I will not speculate on the historical origins of ACC Prediction in Slavic, a topic with a rich literature of its own. What is important here is that the outcome of these events is a synchronic system that might be to some degree ranked but (like the other complex agreement systems discussed by Puljak (1984)) is nevertheless fully consistent with the requirements of universal grammar - which is to say that the system provides an excellent place in which to explore the consequences of particular theoretical hypotheses, such as those concerning component interfaces, covert grammatical categories, and the internal structure of syntactic features like CASE.

References

- Cooper, Robin. (1986). Swedish and the Head Feature Convention. In *Topics in Scandinavian syntax* (ed. by Lars Hellan and K. Koch Christensen). Dordrecht: D. Reidel.
- Crockett, Dina B. (1976). *Agreement in contemporary standard Russian*. Columbus OH: Slavica.
- Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag. (1985). *Generalized phrase structure grammar*. Oxford: Basil Blackwell.
- Klenin, Emily. (1983). *Animacy in Russian: A new interpretation*. Columbus OH: Slavica.
- Maltzoff, Nicholas. (1984). *Essentials of Russian grammar*. Lincolnwood IL: National Textbook Co.
- Moravcsik, Edith A. (1978). Agreement. In *Universals of human language*, volume 4, *Syntax* (ed. by Joseph H. Greenberg), 331-74. Stanford CA: Stanford Univ. Press.
- Perlmutter, David M. and Janez Orešnik. (1973). Language-particular rules and explanation in syntax. In *A Festschrift for Morris Halle* (ed. by Stephen R. Anderson and Paul Kiparsky), 419-59. New York: Holt, Rinehart and Winston.
- Pullum, Geoffrey K. (1984). How complex could an agreement system be? *Eastern States Conference on Linguistics* 1.79-103.
- Zwicky, Arnold M. (1985a). How to describe inflection. *Berkeley Linguistics Society* 11.372-86.
- Zwicky, Arnold M. (1985b). Rules of allomorphy and syntax-phonology interactions. *Journal of Linguistics* 21.2.431-6.
- Zwicky, Arnold M. (1986a). German adjective agreement in GPSG. To appear in *Linguistics* 24.5. Earlier version in Ohio State University Working Papers in Linguistics 31.160-94.
- Zwicky, Arnold M. (1986b). Imposed versus inherent feature specifications, and other multiple feature markings. *Indiana University Linguistics Club Twentieth Anniversary Volume*, 85-106.
- Zwicky, Arnold M. and Geoffrey K. Pullum. (1986). The Principle of Phonology-Free Syntax: Introductory remarks. Ohio State University Working Papers in Linguistics 32.63-91.