Chapter 18

Split-morphology and lexicalist morphosyntax: The case of transpositions

Andrew Spencer

University of Essex

One of Anderson's many contributions to morphological theory is the claim that morphology is split between syntactically mediated inflection and lexically mediated derivation. In Minimalist morphosyntax all morphology is syntax. This means that the split morphology proposal is not meaningful for that model. In lexicalist models, however, the split morphology hypothesis manifests itself as a distinction between direct accessibility to syntactic representations (inflection proper) and lack of accessibility. However, there are construction types which bring the inflection-derivation distinction into question. One of these is the transposition, as illustrated by the ubiquitous deverbal participle. This is a mixed category, being at once a form of a verb yet having the external syntax of an adjective. It is thus unclear which side of the split participles fall. Similarly, participles seem to be an embarrassment for the Word-and-Paradigm models of inflection which have become dominant since Anderson first introduced them to contemporary theorizing. This is because they seem to require us to define a "paradigm-within-a-paradigm" (or "quasi-lexeme-within-a-lexeme").

I provide an analysis of Russian participles within Stump's PFM2 model, deploying the model of lexical representation developed in Spencer (2013), which fractionates representations into more finely grained subcategories than is usual. I take a participle to be the adjectival representation of a verb, coded directly by means of a set-valued feature, REPR. I show how a set of rules can be written which will define the adjectival paradigm as a set of forms belonging to the overall paradigm of the original verb lexeme. The rules define a partially underspecified lexical entry for the participle ("quasi-lexeme"), which has essentially the same shape as the lexical entry for an (uninflected) simplex adjective. Thus, the participle's lexical entry is that of an adjective, just as though we were dealing with derivation, but it realizes the verbal properties of voice/aspect and it shares its semantics and lexemic index with its base verb, just as in the case of verb inflection. The participle thus straddles the split, but in a principled fashion.

Andrew Spencer

1 Introduction: Morphological architecture

Since Anderson (1977) (re-)introduced to generative grammar the traditional notion of "word-and-paradigm", and particularly the ground-breaking work of Matthews (1972) on Latin inflection, morphologists have been grappling with the challenge of providing an adequate characterization of the key notions "word" and "paradigm".

Central to this debate has been the fate of the Bloomfield/Harris interpretation of the morpheme concept (Anderson 2015) and the notion of "Separationism" (Beard 1995). The "word" notion presupposes (at least) a word/phrase distinction. (In morpheme-based approaches no such distinction is necessary and all morphology and syntax can be subsumed under a model of morphotactics.) There are very well known problems with any attempt to find necessary and sufficient conditions for the 'concrete' instantiations of word — the phonological word, (inflected) word form, syntactic word, even. This is generally on account of incomplete grammaticalization, which strands constructions and formatives in a limbo between the status of function word – clitic – affix, compound element– affix, analytical syntactic construction – periphrasis and so on. However, the 'abstract' notions of word are no less problematic, specifically the lexeme and the morphosyntactic word (i.e. an inflected word form together with the morphosyntactic property array that it realizes). Defining the set of morphosyntactic words often requires us to make decisions about what constitutes a word form, which brings us back to the issue of clitics, periphrasis and so on. It also requires us to make sometimes arbitrary decisions about morphosyntactic property sets (MPSs) in the light of form-content mismatches such as (some types of) syncretism (Baerman, Brown & Corbett 2005), overabundance (Thornton 2012), defectiveness (Sims 2015), and deponency (Baerman et al. 2007), as summarized in Stump (2016a).

Inflectional properties, and word-oriented functional categories generally, such as definiteness (for nouns) or modality (for verbs) in English, seem to presuppose an inflection ~ derivation dichotomy that is notoriously hard to pin down. Broadly speaking it distinguishes the creation of new lexical items/units (generally, Saussurean signs pairing a cognitive meaning with a set of forms) from forms of a lexical item/unit. The component of grammar that defines new lexical units or lexemes is derivational morphology. However, as Spencer (2013) itemizes in some detail, such a (canonical) inflection/derivation dichotomy represents just two poles of a scale of types of lexical relatedness. Some of the intermediate types of relatedness pose problems for any clean characterization of the lexeme concept (part of the problem of "lexemic individuation", Spencer 2016b).

One way of characterizing the core of the inflection/derivation distinction is the notion of split morphology (Anderson 1982: and subsequent references). The essence of the distinction can be thought of as an interface claim: inflection interfaces directly with syntax ('inflection is what is relevant to syntax'). The obverse to this claim is that derivation interfaces with the lexicon, in the sense that derivation is what gives rise to expansion of the lexical stock (as well as defining relatedness between already fixed lexical entries), in other words derivation is 'what is relevant for the lexicon'. Anderson implements this architectural claim by saying that it is the rules of syntax themselves which define inflectional morphology.

This move raises the important question of what model of syntax we are presupposing. Most versions of the Minimalist Program presuppose something very close to the model of morphotactics proposed by the American Structuralists: the atoms of representation are morphemes, morphology and syntax are identical (it is therefore a terminological choice whether we think of sentence construction as morphotactics or syntax), and the notions lexeme, word form, inflection, derivation, inflectional paradigm are at best heuristic descriptive terms which cannot be given a coherent definition within the model. The natural syntactic framework for investigating a word-and-paradigm, or rather, lexeme-and-paradigm approach to inflection is, perhaps, a lexicalist, or constraints-based, model (Miller & Sag 1997; Sadler & Spencer 2001; Sadler & Nordlinger 2006). In that case the question of split morphology takes on a somewhat different aspect. Rather than claiming that syntactic rules *construct* inflected forms as such, we must say that inflected forms, compared with derived lexemes, are permitted to interact in specific ways with syntactic representations, or equally that inflected forms bear properties which are visible to syntactic representations and principles.

The obvious way to implement this idea is to say that the abstract characterization of an inflected word includes a morphosyntactic description which overlaps with that of a corresponding syntactic representation. A concrete version of this type of overlap is seen in the form-content mapping, as defined in Stump's notion of paradigm linkage (Spencer & Stump 2013; Stewart & Stump 2007; Stump 2002; 2006; 2016a,b). Stump distinguishes morphological properties, the FORM paradigm, from syntactic properties, the CONTENT paradigm. By default these are homologous, but there are many instances of mismatch. For example, Latin syntax distinguishes singular and plural number and a variety of cases, including dative and ablative, but those two cases are never distinguished morphologically for any lexeme in the plural. On the other hand, Spanish verbs have two distinct subparadigms for the imperfect subjunctive but that distinction is nowhere reflected in the syntax. Other mismatches include deponency and periphrasis. To a limited extent we can say that the form-content paradigm distinction is a reflex of covert split morphology: such a distinction is not definable for derivational morphology.

In lexicalist models, the derivational morphology ~ lexicon interface operates over property sets which don't play a direct role in syntax. The hedge "direct" is important: typically, derivation *is* relevant to syntax, in the sense that it changes a lexeme's morphosyntactic class. More subtly, derivation may make appeal to argument structure realization (witness English Subject Nominalizations, *able*-Adjective formation and so on). But if it is assumed that lexemes have a representation of their word class argument structure and other relevant properties, then such syntactically expressed relations can be defined over lexical representations, as extensively argued in the constraints-based literature (see Wechsler 2014: for a review). This is effectively a statement of the doctrine of lexical integrity, at the abstract level of representation as defined by Ackerman & LeSourd (1997).

The conclusion to be drawn is that morphology interfaced with a constraints-based syntax needs to be split in essentially the way proposed by Anderson, but as an abstract architectural property, which sometimes bears a rather complex relation to concrete mor-

phophonological expression. Inflectional morphology maps to syntactic representations in a way in which derivational morphology is unable to, while derivational morphology serves to define (specific kinds of) lexical relatedness. However, there remain interesting cases of violations of lexical integrity with derivational relations, in which syntax appears to have access to the internal structure of the derived word. This paper will argue that such phenomena require us to extend the scope of the split in morphology in a way that takes the notion of "lexeme" as syntactic atom seriously, and which ultimately provides conceptual motivation for a lexicalist interpretation of split morphology.

A case in point is the class of denominal (relational) adjectives in many languages, which allow one noun to modify another by taking on the morphosyntax of an adjective. In European languages, including English and Russian, such adjectives respect lexical integrity, in the sense that the base noun is opaque to syntax. For example, the base noun *kniga* 'book' in the Russian relational adjective *knižnyj* 'pertaining to a book/books' is opaque to agreement, government or any other syntactic process, just as in English the noun *tide* in *tidal* is opaque. For instance, the phrases *poderžannaja kniga* 'second-hand book' and *knižnyj magazin* 'book shop' do not gives us *poderžannaja/poderžannyj knižnyj magazin*, and although we can say *high tide* and *tidal barrier* we can't say *high tidal barrier*. The importance of these observations is that there are languages in which just such attributive modification into a derived adjective is possible (see the discussion of Tungusic and Samoyedic examples in Nikolaeva (2008), and also the detailed discussion of the Samoyedic language Selkup in Spencer 2013: Chapter 10).

The relational adjectives of Russian and English, however, share one important property with the Tungusic and Samoyedic pure relational adjectives, namely, they have precisely the same lexical semantics (cognitive content) as the base noun. This leaves us with the question of how to explain why in some languages relational adjectives are opaque and in others they are transparent to attributive modification.

Spencer (2013) argues that the crucial difference between true transpositional relational adjectives of, say, Selkup, and the "fake" transpositions of English/Russian is that true transpositions are effectively forms of the base noun lexeme, while the English and Russian relational adjectives are distinct lexemes, though ones which have a semantic representation identical to that of their base, what Spencer (2013: 275) calls a transpositional lexeme. Other types of transpositional lexeme include English property nominalizations (*kindness, sincerity, ...*), deverbal nominalizations such as *destruction*, and participial forms which have been converted into qualitative adjectives such as *(very) boring/bored, charming, excited, ...* A relational adjective which is a true transposition permits inbound attributive modification because it is, in an important sense, still a noun, just as a noun stem marked for number, case, possession or definiteness is still a noun.

One consequence of this reappraisal of the morphology~syntax interface is that the crucial divide can no longer be straightforwardly equated with a traditional distinction between inflection and derivation. It is not appropriate to think of a deverbal participle or a relational adjective as merely an inflected form of a verb or noun, because that participle or adjective will in general inflect like an adjective, not like a verb/noun. However, following Haspelmath (1996), Spencer (2013: Chapter 10) argues for an enrichment

of the traditional notion of the inflection paradigm to include an attribute REPRESEN-TATION (taken from Russian descriptive practice). Thus, a participle is the adjectival representation of the verb, and as such it can have its own adjectival inflectional paradigm. It thus has the outward appearance of an autonomous lexeme, but appearances are deceptive. Rather, the transposition is a "quasi-lexeme", and the transpositional relationship therefore represents a particularly striking instance of a deviation from inflectional canonicity.[1] Like transpositions, these are not usually described under the heading of morphology-syntax mismatches, and like transpositions they are not discussed in Stump's (2016a) otherwise very detailed survey.

In the model of lexical representation argued for in Spencer (2013) the notion "form of a lexeme" in this somewhat extended sense is reflected very simply: all forms of a lexeme share their Lexemic Index. This leads us to propose a (no doubt too strong) hypothesis about the nature of the split in morphology:

Principle of Lexemic Transparency
Let D be a word derived by some regular morphological process from a word B, possibly of different morphosyntactic category. If morphosyntactic processes treat D in the same manner as they would treat the base word, B, even where the category of D is such that we would not otherwise expect it to be subject to such processes, then D is a form of the lexeme B (shares B's Lexemic Index).

I shall argue that the architectural equivalent of splitting inflectional from derivational morphology is this modification of the notion of lexical integrity: if morphology defines a set of forms of a lexeme, rather than defining a new, autonomous lexeme, then those forms will show lexical transparency. Derived lexemes, however, show lexical opacity (one reflex of which is the more familiar property of lexical integrity). This paper will illustrate that proposal on the basis of the behaviour of Russian deverbal participles. These are particularly useful. First, Russian adjectival morphosyntax is very clearly distinguished from noun or verb morphosyntax, so it is easy to show that the participles behave like adjectives. Second, the Russian past tense and conditional mood are expressed by a form which is historically a participle and which show participle-like agreement, but which has been reanalysed as a verb form (the l-participle). This contrasts in important ways with the true participles. Third, like many languages, Russian often converts its participles into true qualitative adjectives. These have (almost) exactly the same set of forms as the participles but their syntax is no different from that of a simplex adjective. The true participles are like the l-participle in showing lexical transparency with respect to the base verb. They therefore both appear on the inflectional side of the split morphology. This is because they are forms of the verb's paradigm, and do not constitute independent lexemes in their own right. They contrast with the converted participles,

---

[1] Other such deviations are certain forms of evaluative morphology (cf "the diminutive *form of* a noun") and grammaticalized argument structure alternations (cf "the passive/anti-passive/applied/causative *form of* a verb").

which are autonomous lexemes and hence opaque with respect to the verb properties of their (etymological) base lexeme.

2 Lexical representations and lexical relatedness

2.1 Introduction

Our discussion will require us to be explicit about a number of aspects of lexical representation and the way that words, in the broadest senses of the term, are related to each other. I shall adopt a generalized form of Stump's (2001) Paradigm Function Morphology, which I have called *Generalized Paradigm Function Morphology*, GPFM (Spencer 2013). The GPFM model is designed to permit us to use the machinery of PFM to describe types of lexical relatedness which go beyond normal inflectional morphology. It thus extends the lexical representations that morphology has access to by incorporating representations of syntactic properties and the lexical semantic representation of words. In GPFM lexemes have to be individuated by means of an arbitrary index, the Lexemic Index (defining something like the key field in a database). One of the reasons for this is because it is arguably not possible in the general case to individuate lexical representations of lexemes in terms of any of the linguistically relevant properties that can be ascribed to a lexical representation. In addition, however, the index serves an important role in distinguishing certain types of morphological relatedness.

2.2 Lexical representations

I begin with the descriptive representational apparatus required to characterize an inflected word form, taking inflection to be an uncontroversial category for the sake of exposition. I then generalize the representational format to provide a characterization of the lexemic entry.

A word has a minimum of three contentive attributes (together with a fourth, its Lexemic Index, LI): FORM, SYN(TAX), SEM(ANTICS). The SYN attribute records idiosyncratic selectional or collocation properties, but its main component is the argument structure attribute, ARG-STR. This records thematic argument arrays in the standard fashion (notated here as x, y, … variables). However, it also includes a semantic function (sf) role.

For nouns and verbs the sf roles are the "R" and "E" roles respectively, familiar from the literature. The "R" (for "referential") argument is canonically associated with lexical entries whose SEM value belongs to the ontological class of *Thing*. It thus identifies those predicates that typically denote (concrete or abstract) objects and that can serve as the lexical head of a referring expression, i.e. a canonical noun. Thus, the "R" argument of TREE[2] corresponds to the "x" variable in the semantic representation $\lambda x.\text{TREE}(x)$. It is the argument that is the target of attributive modifiers: *(tall) tree* (Spencer 1999). See Lieber (2004: 16; 55–59) for concrete examples of the R role being deployed in morphology. The

---

[2] Where relevant, I adopt the standard convention of putting the name of a lexeme in SMALL CAPS.

"E" (for 'event(uality)') argument (sometimes written as "e" or "s" (for "situation")) is canonically associated with lexical entries whose SEM value belongs to the ontological class of *Event*. It thus identifies those predicates that typically denote states or events (eventualities) and that can serve as the lexical head of a clause, i.e. a canonical verb. Thus, the "E" argument of FALL corresponds to the "e" variable in the (neo-Davidsonian) semantic representation $\lambda e, x.FALL(e,x)$. For attributive modification (principal role of the traditional adjective class) I assume a semantic function role labelled "A". This is coindexed to a noun's R sf role to represent attributive modification. All adjectives which function as attributive modifiers, including relational adjectives and participles, have an "A" semantic function role.

I assume the SEM attribute is essentially a formula in predicate calculus defined over the ontological types *Thing, Event, Property* (Jackendoff 1990), and perhaps others, corresponding loosely to the morphosyntactic categories of Noun, Verb, Adjective. I remain here agnostic as to whether the categories N, V, A are universal and if so, in what sense. I assume that some languages also have a category of Adverb, and also transpositional morphosyntax, adjective-to-adverb (as in English *ly*-suffixation), verb-to-adverb (gerund) and noun-to-adverb (found in Selkup, for example), but I will not have much to say about that category here.

Adpositions may mandate a further ontological category of, say, *Relation*, but I ignore that too. The SEM attribute can be thought of as a label for an encyclopaedic representation, such as $\lambda x.CAT(x)$ or $\lambda x,y.WRITE(x,y)$, but sometimes including linguistically encoded information relevant to semantic interpretation that cannot simply be consigned to an undifferentiated encyclopedia, for instance, $\lambda x,y,\delta.SIMILAR\_TO(x,y,\delta)$ $\land CAT(y) \land DIMENSION(\delta)$, 'similar to the property of "cat" in some dimension, δ', or $\lambda x,y.AGAIN(WRITE(x,y))$ 'to re-write something'.

The FORM attribute is essentially a record of the word's morphology. Assuming an articulated inflectional system, complete with arbitrary inflectional classes and possibly other purely morphological, paradigm-based properties, the FORM attribute needs to specify all the information needed to locate the word form in the appropriate inflectional paradigm. The first property is the morpholexical category, MORCAT. This will typically be derived by default from the syntactic category of the representation (the SYN|CAT attribute),[3] but that default mapping is not infrequently overridden, sometimes in rather complex ways.

The next property is largely defined by reference to the syntactic category of the word form/lexeme, namely, the morpholexical signature, MORSIG. This specifies all those morphosyntactic properties for which an element of that MORCAT is obligatorily inflected. An inflected word has to have a specification of the morphosyntactic property set (or sets), MPSs, that it realizes. In the case of syncretism this may be a (natural or unnatural) class of MPSs. For example, a Russian adjective is obligatorily inflected for at least

---

[3] Given the complexities of category mixing it is better to dispense entirely with morphological or syntactic category labels such as "verb", "adjective". The required lexical classes can be defined over other aspects of representation much more efficiently and it is not difficult in constraints-based models to ensure that those aspects of representation are accessible to rules of morphosyntax. However, for convenience of exposition I will continue to talk of (morphological or syntactic) verbs, adjectives and so on.

the properties of number, gender and case, and these features are therefore listed in the MORSIG (a gradable adjective is also inflected for comparative and superlative forms). We will see in §5 that the conception of MORSIG assumed in Spencer (2013) can be enriched and extended to include the FORM-CONTENT paradigm distinction introduced in Stump (2006) and subsequent work.

Finally, the representation has to specify a phonological form for the word, through an attribute FORM|PHON. The precise characterization of the PHON entry, in general, will be given by the rules of inflectional morphology. I will assume that one aspect of the PHON representation will be a specification of the STEM on which the inflected form is based, but I ignore this refinement because it will not be relevant to the question of split morphology.

The actual inflected word forms of the lexeme are defined by inflectional rules, which apply to the pairing $\langle \text{£}, \sigma \rangle$, where σ is a complete, permissible feature set for the lexeme with Lexemic Index £. The lexemic representation has to include all the idiosyncratic morphological information relevant to a lexeme's realized paradigm. In the next subsection I summarize the way that the Paradigm Function can be generalized to define not only inflection but all the systematic forms of relatedness.

One aspect of these representations is worth noting. In keeping with the inferential-realizational assumptions underlying our model of inflection the SEM attribute remains constant for all inflected forms. In particular, there is no characterization at the level of the lexical representation of a word form (much less the level of lexemic representation) of the semantics of, say, tense or number. What this means is that in the syntax the VP which is headed by a past tense form verb may, ceteris paribus, be interpreted as referring to an event situated prior to speech time. However, since 'past tense' forms are also used in sequence of tense constructions, irrealis conditionals and so on, 'past time' is only the default interpretation.

2.3 Lexical relatedness

We can now ask what types of systematic relatedness lexemic entries (i.e. lexemes) can exhibit. Spencer (2013) argues extensively that we can find pretty well all the logically possible types as defined by the very crude but simple artifice of defining non-trivial differences in the four principal attributes, FORM, SYN, SEM, LI. For instance, if we consider pairs of representations of distinct lexemic entries, $\text{£}_1, \text{£}_2$, i.e. those with distinct LIs, then we can identify several different types of relatedness (usually all treated as derivation).

Suppose that the lexemes $\text{£}_1, \text{£}_2$ are distinct in FORM, SYN, SEM representations. Suppose also that the FORM/SEM representations of £_2 subsume or properly include (in some sense) those of £_1. Then we have standard (canonical) derivational morphology, DRIVE $\Rightarrow$ DRIVER. Languages sometimes define derived lexemes without changing the FORM attribute at all, however. A case in point is that of adjectives which are converted wholesale to nouns without any change in morphology (Spencer 2002; see also the discussion of *Angestellte(r)* nouns in Spencer 2013). We will later see examples of derivation

in which FORM/SYN/LI attributes are changed but without changing the meaning (transpositional lexeme).

Now let's consider what happens if we keep the LI constant, that is, we consider intra-lexemic relatedness. To begin with, let us assume that only the FORM attribute can change. In the canonical case this is the same as inflectional morphology. Ignoring for the present the transpositional interpretation of the participles as a verbal noun or as verbal adjectives, we can say that all inflected forms of SING are forms of a verb. We have to be a little cautious when referring to syntactic properties: the syntactic distribution of any given inflected form is, in general, distinct from that of other forms. The 3sg subject agreement form of a verb does not occur in the same syntactic positions as the 3pl form. The properties that give rise to these differences, however, are precisely the MPSs which bifurcate into FORM/CONTENT paradigms. This means that we must enrich lexical representations in the obvious way: FORM paradigms are defined over features typed as FORM features, and CONTENT paradigms are defined over features typed as SYN features, with the proviso that the two sets of features are identical by default. Modulo the CONTENT paradigm, then, in canonical inflection the SYN value of a given word form is identical to that of the other forms of that lexeme. This means that most inflection is what Booij (1994) would call contextual inflection.[4]

Likewise, the lexeme SING denotes the same event type in all of its inflected forms, and in that sense all word forms share the same SEM representation. In Spencer (2013) I argue that there are certain types of inflection that enrich the semantic representation of the base lexeme, whilst remaining inflectional. Certain kinds of Aktionsart marking, as well as semantic case marking often have this characteristic, as do causative argument structure alternations. In traditional descriptions of languages with such inflection we often find terminological vacillation, as linguists are unsure whether to label, say, the iterative form or the causative of a verb inflectional or derivational (and similar problems afflict evaluative morphology).

The GPFM descriptive framework proposed in Spencer (2013) generalizes the PFM model so that all forms of lexical relatedness, from contextual inflection to derivation, are defined over four principal attributes of a lexical representation. This requires us to generalize the notion of the Paradigm Function to that of a Generalized Paradigm Function (GPF), which is like the Paradigm Function except that it consists of four component functions, $f_{form}, f_{syn}, f_{sem}, f_{li}$. For canonical derivation the GPF introduces non-trivial changes to all four components (including the LI). For the converted adjectives and *Angestellte(r)* nouns the f_{form} function has no effect (it can be thought of as a kind of identity function). For most inflection, the $f_{syn,sem,li}$ functions are the identity function, because the GPF simply realizes inflectional properties of the lexeme at the FORM level.

In the GPFM model the lexemic representation is defined in terms of the Lexemic Index and a completely underspecified (empty) feature set, u, for example, $\langle \text{PUT}, u \rangle$, a special case of the GPF. This maximally underspecified GPF defines just those properties of a

---

[4] This includes Booij's parade examples of inherent inflection, past tense and plural number. See Spencer (2013: 77–82) for critical discussion of Booij's distinction.

lexeme (identified by its Lexemic Index) that are completely idiosyncratic and completely unpredictable. However, although such a representation reflects the traditional notion of a maximally compact, non-redundant dictionary entry, it is not a representation that can serve as the direct input to rules of inflection. This is because the lexemic entry has to be specified for those inflectional properties that it can and must inflect for. This set is defined by the morphological signature, MORSIG. In Spencer (2013: 199) I make this rather obvious point explicit in the *Inflectional Specifiability Principle*. In the current context this can be stated as follows: a lexeme is inflected for a given MPS, iff that MPS is defined in its MORSIG.

In Spencer (2013) I treat the MORSIG attribute as part of the FORM paradigm of a lexeme. However, we know that the FORM and CONTENT paradigms of a lexeme can differ substantially. For this reason, it is necessary to enrich the SYN attribute of a lexemic entry with a (possibly distinct) MORSIG attribute. The values of the (FORM and CONTENT) MORSIG attribute are for the most part predictable from other aspects of the lexical representation. First, the FORM MORSIG attribute is generally projected from the CONTENT MORSIG and by default the two attributes are identical. Second, to some extent the meaning of the lexeme can determine the content of the MORSIG attribute. Most importantly, however, the MORSIG (which, recall, is essentially a record of the properties for which a lexeme inflects) is largely projectable from various aspects of the SYN attribute, notably the ARG-STR attribute. Thus, a lexeme with the SYN|ARG-STR value $\langle E \langle x, \ldots \rangle \rangle$ (i.e. a verb) will by default have the syntax and morphology of a verb. The lexemic representation needs to be enriched to include purely idiosyncratic information, such as irregular stem forms, irregular inflections, defective cells or subparadigms, and so on. Technically, this can easily be achieved in GPFM by defining very specific functions for particular properties over the LIs of the lexemes concerned. For instance, the irregular past tense of PUT can be defined by a function defining the $STEM_{pst}$ form for the pairing $\langle \text{PUT}, u \rangle$: GPF($\langle \text{PUT}, \{STEM_{pst}|PHON\}\rangle$) = /pʊt/ or similar. This will override any less specific (in practice, any other) statement of past tense morphology. Similarly, a defective lexeme such as FORGO (lacking a past tense form) will have a tense-specific GPF($\langle \text{FORGO}, \{tense:pst\}\rangle$) = undefined. This again will override any other statement, including the GPF($\langle \text{GO}, \{tense:pst\}\rangle$) = /wɛnt/, which applies to one other verb based on GO (cf *underwent*) and hence is less specific.

The role of the MORSIG attribute can be simply illustrated by the English plural. Any lexeme with the SYN|ARG-STR|$\langle R \rangle$ value licenses MORSIG|num:{sg, pl}, provided that its SEM attribute specifies it as a count noun. For a mass noun the MORSIG value will be just num:{sg}, while for a plurale tantum noun it will be num:{pl}. The {pl}, resp. {sg} values for such nouns are therefore undefined, so that a GPF($\langle \text{SINCERITY}, \{num:pl\}\rangle$) or GPF($\langle \text{SCISSORS}, \{num:sg\}\rangle$) will not correspond to any legal output.

In summary, I assume a representation for a dictionary entry of a very traditional kind: it is minimally redundant, specifying just the unpredictable phonological and semantic information, together with any morphosyntactic information that cannot be projected from the phonological and semantic specifications. For an entirely well-behaved lexeme

belonging to the default inflection class for that word category, this is all the information that is required, but that is only because the morphological signature can be projected from that entry too.

To all intents and purposes GPF collapses with PFM for most cases of inflectional morphology. However, the model is designed to cover all types of lexical relatedness, up to regular derivation, but using essentially the same machinery as PFM. For instance, for the derivation of a Subject Nominal such as DRIVER from DRIVE we would have the partial GPF shown in (1), where $\mathcal{V}$ is the Lexemic Index of a verb lexeme and δ is the derivational feature which defines the Subject Nominal formation process.

(1) a. $f_{li}(\langle \mathcal{V}, \delta \rangle) = \Delta(\mathcal{V})$, where Δ is a function over LIs corresponding to the derivational feature δ

   b. $f_{sem}(\langle \mathcal{V}, \delta \rangle) = [\textit{Thing}\ \lambda x, \text{PERSON}(x) \wedge \mathcal{P}'(x)]$, where $\mathcal{P}'$ is a suitable form of the semantic representation of the lexeme $\mathcal{V}$

   c. $f_{syn}(\langle \mathcal{V}, \delta \rangle) = u$

   d. $f_{form}(\langle \mathcal{V}, \delta \rangle) = Z \oplus er$, where Z is the δ-selected stem form of $\mathcal{V}$

This corresponds to a novel lexemic entry which is exactly like that in (1) except that it is defined over the pairing $\langle \Delta(\mathcal{V}), u \rangle$. If DRIVE is the LI of the verb DRIVE, then $\Delta(\mathcal{V})$ is the LI of the derived lexeme DRIVER and $\langle \Delta(\mathcal{V}), u \rangle$ defines its lexemic entry.

The representation in (1) lacks any syntactic specification. This is because that specification can be given by a default mapping from the SEM attribute, by virtue of the Default Cascade (Spencer 2013: 191–194). Under that principle, regularly derived lexemes have their principal morphosyntactic properties projected from their semantic representations, in accordance with the notional model of parts of speech. Thus, DRIVER is of ontological category *Thing* and therefore by default has the semantic function role R. I assume that the SEM attribute includes an indication that the lexeme denotes a countable *Thing* so that the lexeme licenses the full MORSIG|NUM:{sg, pl}. However, the lexeme in (1) is derived, not simplex. In Spencer (2013) I propose a Category Erasure Principle, under which the morphosyntactic properties of a base lexeme in derivation are deleted so that they can be overwritten by the Default Cascade. However, given maximally underspecified lexemic entries to start with this is probably not necessary: the word formation rule interpretation of the GPF takes the only information it has available in a lexical entry, that is the phonology of the root and the meaning, and modifies it, say, by adding an affix, and by enriching the SEM representation systematically, for instance, by adding a predicate. The Default Cascade then specifies the underspecified properties, including the two MORSIG representations. If the base lexeme's entry includes non-default specifications such as irregular inflected forms or non-standard argument realization, these will override the Default Cascade. Illustrative examples of simple inflection and derivation are provided in the Appendix.

I have sketched the GPFM approach to inflection on the one hand, and standard derivation on the other hand. Non-standard types of derivation are handled by overriding

some of the defaults in the GPF. However, this still leaves us with the somewhat intriguing, but very widespread set of relatedness types in which FORM/SYN attributes are altered, much as in derivation, but without changing either the SEM representation or, crucially, the LI attribute, one of the class of relatedness types often called a "transposition". The parade example of this type is, perhaps, the deverbal participle, the 'adjectival representation' of a verb.

Transpositions pose particular difficulties for a simple interpretation of the inflection-derivation divide (and, indeed, for any architecture with the equivalent of split morphology): on the one hand, a participle is generally taken to be 'a form of' the base verb (regular participles are never given their own entry in a traditional dictionary, for instance). On the other hand, a participle is morphologically and syntactically a kind of adjective. If morphology is split, on which side do transpositions fall?

3 Paradigm Linkage: the problem of transpositions

Spencer & Stump (2013) (see also Bonami & Stump 2016; Stump 2016a) describe a variant of the Paradigm Function Morphology model, PFM2, which explicitly distinguishes two types of paradigm. FORM paradigms determine the mapping between MPS's and the word forms realizing them; CONTENT paradigms determine the set of grammatical distinctions a lexeme needs to make in a syntactic representation. The two paradigms are related by rules of paradigm linkage. In the default case the two paradigms align perfectly: a morphological plural noun is syntactically plural and vice versa. However, there are numerous mismatches. A simple case in point is provided by English perfect and passive participles. These have distinct syntax (they collocate with different auxiliaries, and only the passive participle can be used attributively), but they are never distinguished morphologically. Thus, the (single) FORM property (say, VFORM:en-ptcp, or the output of a function f_{en}, as in Aronoff 1994) has to map to two CONTENT properties. Stump (2016a) provides an extensive survey of the principal types of mismatch encountered in the world's inflectional systems.

I will argue that we can regard the FORM/CONTENT paradigm distinction as a reflex of split morphology for constraints-based lexicalist models of syntax. The reasoning is very simple — no corresponding FORM/CONTENT distinction can be mandated for derivation. Indeed, given standard PFM2 assumptions it is difficult to imagine what such a distinction could mean.

Incorporating the FORM-CONTENT paradigm distinction into GPFM has consequences for the way in which lexical representations are organized. We have seen that the FORM attribute of a lexical representation is defined in part in terms of its MORSIG, which determines precisely those properties a lexeme must inflect for. In the original GPFM model the MORSIG attribute is only defined at the level of FORM paradigms, therefore. Since CONTENT paradigms are not always congruent to FORM paradigms we will need to draw the appropriate distinction in lexical representations. The obvious

way to do this is to assume that each lexeme is associated with *two* (at least) MORSIG attributes, one of which is a value of the FORM attribute and the other of which is a value of the SYN attribute. We can call these f-MORSIG, c-MORSIG respectively.[5]

Stump (2016a: 253) defines paradigm linkage in terms of the Paradigm Function and a ***Corr*** function. Given a lexemic index £, a set of MPSs σ, τ, and a stem form, Z, the function ***Corr***($\langle$ £,σ$\rangle$) delivers a form correspondent $\langle$Z,τ$\rangle$, such that if PF($\langle$Z,τ$\rangle$) = $\langle$w,τ$\rangle$, then PF($\langle$£,σ$\rangle$) = $\langle$w,τ$\rangle$. By default the FORM and CONTENT features are identical, that is the set σ = τ, as defined by a set of property mappings (***pm***), that is, ***pm***(σ) = σ. However, the function ***pm*** is called upon whenever there is a mismatch between FORM and CONTENT properties. This is the case, for instance, with deviations such as syncretism, deponency and so on. In the case of active~passive deponency the property mapping maps the morphological passive voice forms to the syntactic active paradigm and leaves the morphological active voice forms undefined.

We can ask how paradigm linkage would work for derivation, by taking the example of a derivational feature, δ, such as the privative denominal adjective feature, *privadj*, which derives *friendless* from *friend*, as described in Stump (2001:252–60).[6] Here, we are dealing with trivial (two-celled) paradigms, so that the featural mismatches of inflectional paradigms will not be found, and we can work with just a single derivational feature, δ.

(2) Let ***Corr***($\langle$FRIEND, *privadj*$\rangle$) be the correspondence function for *privadj* applied to the lexeme FRIEND.
Corr($\langle$FRIEND, *privadj*$\rangle$) = $\langle$Z,δ$\rangle$, where Z = |friend|, the stem of FRIEND, and δ = *privadj*. If PF($\langle$Z,δ$\rangle$) = $\langle$w,δ$\rangle$, then PF($\langle$£,δ$\rangle$) = $\langle$w,δ$\rangle$, hence, PF($\langle$FRIEND, *privadj*$\rangle$) = $\langle$friendless, *privadj*$\rangle$.

This seems very straightforward, but there is a hidden difficulty, not immediately apparent from a language like English, with limited inflection. Consider a hypothetical language just like English but in which nouns and adjectives have entirely different inflectional paradigms, or, indeed, consider a derivational process such as that of Subject Nominalization which derives DRIVER from DRIVE. What we have to ensure is that the output lexeme is (automatically, by default) inflected as a noun, as opposed to the base lexeme which is inflected as a verb. As it stands, the Paradigm Function applied to a pairing of Lexemic Index and derivational feature will not deliver what Stump calls the realized paradigm of the output lexeme. At best, it might define the (or an) uninflected stem form of the derived word. Moreover, without significantly altering the nature of the ***Corr*** function it will be impossible to define additional lexical information such as

---

[5] A number of authors have proposed that cells in a lexeme's paradigm can be realized by multiword, periphrastic, constructions (Sadler & Spencer 2001; Brown et al. 2012; Bonami 2015). Popova & Spencer (forthcoming), following suggestions by Bonami (2015), propose that such constructions demand additional CONTENT feature sets specifically to define the content of the periphrastic expression, which is often at odds with the default feature content of the words which make up that expression. Periphrasis therefore provides substantial motivation for a FORM/CONTENT or m-/s-feature (Sadler & Spencer 2001) distinction.

[6] Stump does not discuss derivation, or, indeed, transpositions in his works on paradigm linkage.

inflectional class membership, idiosyncratic stem forms or other deviations from default inflection, or indeed any non-default or purely morphological property of the output.[7]

A direct solution to this problem might be to enrich the content of the derivational feature, δ, so that it incorporates the full set of paradigmatic oppositions realizable by the derived lexeme. Thus, the Subject Nominal feature (say *subjnom*) could be defined as the complex [*subjnom*, NUMBER:α]. This would mean that we would have to define the Paradigm Function so that it defines each inflected form of the output lexeme, rather than defining an underspecified lexemic entry, which then gets inflected by the standard inflectional rules applying to words of that category. For instance, the Paradigm Function could take the form PF(⟨DRIVE,{*subjnom*, NUMBER:sg}⟩) = ⟨driver, {*subjnom*, NUMBER:sg}⟩, PF(⟨DRIVE,{*subjnom*, NUMBER:pl}⟩) = ⟨drivers, {*subjnom*, NUMBER:pl}⟩. We could call this approach the 'full-listing approach'.

Now, the full-listing approach would have the rather peculiar consequence that the word forms {*driver, drivers*} would be inflected forms of the verb DRIVE, and there would be no such thing as a DRIVER lexeme. Not only is this entirely counter-intuitive, and at variance with any sensible distinction between inflection and derivation, it becomes even more counter-intuitive when we see recursive derivation. It would entail that the noun *reprivatizability* was an inflected form of the adjective *private*. The problem is that the ***Corr*** function needs to be able to define not a cell in a realized paradigm (what syntacticians refer to, misleadingly, as the lexical entry of a word(form)), but it has to define a featurally underspecified lexemic entry (an autonomous dictionary entry in traditional terms). I therefore reject the full-listing approach in favour of that adopted in the GPFM model.

We are now in a position to examine the case of transpositions. The problem we must address is how to ensure that the transposition is assigned to its own inflectional paradigm, proper to its new morphosyntactic category, whilst still in some sense remaining part of the inflectional paradigm of the base lexeme. Recall that I have proposed adopting a class of features, [REPRESENTATION:{...}] to define the paradigm space of a transposition. I now consider the way this feature can be deployed to define the inflectional paradigm-within-a-paradigm of a transposition. Let ρ = [REPR:κ] for some transpositional relation κ, e.g. a verb-to-adjective (participle) transposition. I assume that the GPF applied to the pairing ⟨£,ρ⟩ defines a partially specified representation for the transposition. Normally, when the Paradigm Function applies to a pairing of LI and MPS the MPS has to represent a complete and coherent set of features sufficient to define the inflected form. However, this is not strictly speaking a property of the PFM system itself. In principle, we could define a partial paradigm for a lexeme by reference to just a proper subset of the features required to define any fully inflected word form. For instance, suppose that verbs in a language inflect for a variety of tense-aspect-mood-voice (TAMV) series, and that in addition they show subject agreement. We could, in principle, define the stem sets for the TAMV categories independently of the agreement morphology, by simply not specifying the subject agreement properties, effectively defining a set of "screeves" for the language (as in the Georgian descriptive tradition, Anderson

---

[7] It is also not clear how the additional semantic predicate of the output would be defined.

1992: 141–45). This makes use of the same notion of feature underspecification used in GPFM to define derivation, but relativized to a specific feature set.

The GPF for a transposition has to specify the derived morphosyntactic category, κ, and, by default, the f-/c-MORSIG attribute, associated with that category, together with additional morphological properties inherited from the base, such as TAMV properties. If the Paradigm Function were to apply in the same way for transpositions as for other inflected forms then the **Corr** function would have to provide the form correspondents for each inflected form, but it would not be able to provide a lexemic entry for the uninflected transposition. In other words, standard application of PFM2 principles would give rise to a full listing interpretation of the paradigm. It would not be possible to provide any characterization of the transposition as a quasi-lexeme (see p. 389). This generates many of the same conceptual and technical problems as the application of **Corr** to derivation. One additional consequence is that it would be difficult to describe the very common situation in diachronic change in which a participle is reanalysed as a qualitative adjective, often without significantly changing the lexical semantics of the original verb lexeme (transpositional lexeme). If we treat the participle as akin to a lexeme, then we can easily define the diachronic reanalysis over that representation, just as we would for ordinary derivational conversion. One of the problems with the full-listing approach is that it is not clear how we could permit the transposition to inherit MORSIG properties of the derived category (Adjective in the case of participles). We therefore need to define the GPF in such a way that it allows us to define a quasi-lexeme as part of the paradigm of the base.

First note that application of the Default Cascade to define derived categories is entirely excluded in the case of (pure) transpositions, since these preserve the meaning, and hence the ontological category, of their base; indeed, this is the whole point of a transposition. The most natural assumption is that the transpositional morphosyntax effects a shift in the syntactic categorization and that morphological recategorization falls out as a consequence. (In fact, the extent to which the transposition acquires derived categorial properties and loses those of its base is subject to a good deal of cross-linguistic variation, tempered by poorly understood typological tendencies. See Malchukov (2004) for discussion in the context of action nominal transpositions.) This is effectively a weaker instantiation of the Default Cascade. Following Spencer (1999; 2013), I will assume that the shift in syntactic representation is actually a modification of the argument structure representation; specifically, the definition of a complex sf role (see below). The crucial point is that the transpositional GPF changes the SYNCAT value of the base lexeme to that of a mixed category and this, ceteris paribus, will automatically entrain a shift in the morphological category and hence the MORSIG attributes.

The basic machinery is only hinted at in Spencer (2013: Chapter ten). In order to develop an explicit account we first need to refine the definition of the REPR(ESENTATION) attribute. I will therefore assume that the REPR feature takes an ordered pair as value, not a singleton element, [REPR:⟨κ,λ⟩], where κ, λ range over lexical categories. Thus, a participle represents a verb as an adjective and hence bears the specification [REPR:⟨V,A⟩], while a predicatively used noun will have the specification [REPR:⟨N,V⟩]. We then cou-

ple these feature values to syntactic category specifications in the obvious way, stated informally in (3).

(3)   Given [REPR:$\langle\kappa,\lambda\rangle$], for λ = N, V, A.
      Then SYNCAT = λ,
      and by default, MORCAT = λ

In the GPFM framework the [REPR] attribute will be associated with an appropriate enrichment of the semantic function role of the ARG-STR attribute. A verb base ARG-STR includes E, the event semantic function role, $\langle E, \langle x,...\rangle\rangle$. That of the participle is enriched by addition of the A semantic function role, to become (simplifying somewhat) $\langle A_i\langle E, \langle x_i,...\rangle\rangle\rangle$, where the co-indexing indicates that the element of which the participle is predicated is an element of the argument array of the base verb lexeme (the highest such argument for Indo-European type participles, though not necessarily so in other languages). This representation reflects the mixed categorial nature of the participle. While its main external syntax (Haspelmath 1996) is now that of an adjective, it retains the eventive ARG-STR of the base verb, which permits it, on a language specific basis, to realize a number of verb properties. These typically include the realization of internal arguments as verb dependents (complete with quirky case marking). The E semantic function role also permits modification by adverbials targeting event semantics.

Exactly which adjectival properties are acquired and, more crucially for participles, which verb properties are lost or retained, differs cross-linguistically. The Indo-European participle, for instance, typically retains the active∼passive alternation, but may also retain aspect (Slavic) or tense (Sanskrit; Lowe 2015). As attributive modifiers, the Indo-European participles can only modify a noun which expresses the participle's highest (subject) argument, effectively making them into heads of subject-oriented relative clauses, but in other languages there is much greater freedom in the choice of argument that can be relativized on by the participle (for discussion see Spencer 2016c).

In §5 'Paradigm linkage rules for Russian' I present a more detailed analysis of Russian participles in which this basic schema for transpositions is expanded upon.

4 Russian participles

In this section we look at the set of four participles that are regularly associated with Russian verbs. Before we can consider these, however, we need to understand Russian verb inflection and the place of the participles in that system. I begin with an overview of the grammatical distinctions as a whole made by verbs, in other words, the CONTENT paradigm, before considering the actual morphological forms themselves. We encounter the familiar problem that there is no consensus on just what the oppositions are and how they relate to each other, and so some of my descriptive decisions will be motivated in part by expositional convenience. I illustrate with the second conjugation transitive verb UDAR′IT′ 'hit', whose imperfective aspect series is formed by shifting to the first conjugation, UDAR′AT′.

The simplest categories are the infinitive and imperfective ("present") and perfective ("past") gerunds. These are indeclinable. Telic verbs such as UDAR'IT'/UDAR'AT', 'hit', require imperfective and perfective aspect forms for most of their paradigm. We can distinguish three moods, indicative, imperative, conditional. The imperative is straightforward and I will ignore it here. The indicative distinguishes three tenses, present (for imperfectives only), past, future. I return to the conditional below.

Transitive verbs show an active~passive voice opposition. However, the voice system is complicated by the fact that imperfective verbs are able to take the reflexive suffix *-s'a/s'*. This has the basic function of giving a reflexive/reciprocal meaning (though similar meanings can also be expressed with fully-fledged reflexive/reciprocal pronouns). The reflexive form also has a wide variety of other uses, including the passive (Gerritsen 1990). This means that we should set up a reflexive voice category and define the imperfective passive in terms of this, but I set this task aside since it is not directly relevant. The perfective verbs express the passive alternation periphrastically, with BE + perfective passive participle. At the level of the CONTENT paradigm we need to be able to define [VOICE:{act,pass}] for both perfective and imperfective verbs, therefore.

Present tense is expressed morphologically, but only for imperfective verbs. There is no dedicated tense marker, and in effect, present tense is realized by the person/number subject agreement morphology. The future tense is expressed periphrastically by imperfective verbs, by means of BE + the infinitive form. For perfective verbs the future is expressed by a paradigm which is essentially the same as the present tense paradigm for imperfective verbs. Thus, the present tense of the imperfective verb P'ISAT' 'write' is *p'išu, p'išeš,...* and the future tense of the prefixed perfective verb form NAP'ISAT' is *nap'išu, nap'išeš,....*

One of the main challenges of the Russian verb system is the representation of the past tense. This is derived historically from a periphrastic perfect tense series formed by BE and a resultative participle expressed by a suffix *-l*, the l-participle. Syntactically, the l-participle behaved like a predicative adjective, agreeing with the subject in number and gender, but not in person. The auxiliary BE was lost, leaving the l-participle and its adjective-like agreement as the sole exponent of past tense. The agreement inflections on an l-participle such as *(na)p'isal* 'wrote' are almost identical to those on a predicative (short-form) adjective such as MAL 'short': *(na)p'isal, (na)p'isala, (na)p'isalo, (na)pisal'i* vs *mal, mala, malo, maly*. The only real difference is that in the plural the l-participle stem is palatalized but not the predicative adjective stem: *(na)p'isal'i* vs *maly*.

The simple way of analysing the past tense construction would be to take the *-l* formative to be an exponent of past tense and define two distinct sets of subject agreement rules for present/future and past tenses. However, the l-participle has one other significant usage which precludes this direct analysis. The conditional mood is expressed by means of the invariable particle *by*, a kind of freely distributed enclitic (it can occur anywhere in the clause except absolute initial position). This can co-occur with the infinitive, (4).

(4) Esl′i by skazat′ pravdu, …
     if   BY say.INF truth

     'To be honest, …'

However, it is much more often found with the l-participle, (5).

(5) Esl′i by ty   skaza-l   pravdu, …
     if   BY you say-L-PTCP truth

     'If you told/were to tell/had told the truth, …'

As is indicated in the gloss, the conditional is tenseless, and serves as the translation equivalent of past or non-past conditional in English. The existence of the conditional construction means that we cannot regard the l-participle simply as the exponent of past tense. In fact, it is an instance of what, since Aronoff (1994) morphologists have called a "morphome", that is, a pure morphological form, which serves as a stem for building up inflected word forms, but which realizes no MPSs on its own and whose distribution is not mandated by any semantic, phonological or other non-morphological properties. The CONTENT paradigm for Russian is summarized in Table 1, but ignoring the true participles.

Table 1: Russian verb CONTENT paradigm for UDAR′IT′/UDAR′AT′ 'hit'

ASPECT	imperfective	perfective
INFINITIVE	udar′a-t′	udar′i-t′
GERUND	udar′a-ja	udar′i-v(ši)
IMPERATIVE	udar′a-j(te)!	udar′(te)!
TENSE		
present	udar′a-ju, -eš, …	\<none\>
future	bud-u, -eš, …udar′at′	udar′-u, -iš
past	udar′a-l, -a, -o, -′i	udar′i-l, -a, -o, -′i
CONDITIONAL	udar′a-l, -a, -o, -′i + by	udar′i-l, -a, -o, -′i + by
PASSIVE	udar′at′-s′a, etc	(byl) udaren, -a, -o, -y

I turn now to the morphological or FORM paradigm. We can divide Russian verb forms into five groups. The first is the infinitive and the second the set of two indeclinable gerund forms. The third is the set of finite forms, including the imperative mood. In practice, these are limited to the present (non-past) forms showing subject agreement in person/number. The fourth is the l-participle. Finally, we have the set of (declinable) active and passive participles discussed earlier. These are tabulated in Table 2.

Summarizing Table 2, the verb can/must inflect for aspect. The verb system also shows voice alternations. However, these are effected either through reflexive morphology (imperfective verbs) or periphrastically (perfective verbs) so voice proper lacks a dedicated, purely morphological exponent, except for the true participles. The infinitive, gerunds

Table 2: Russian verb FORM paradigm for ᴜᴅᴀʀ′ɪᴛ′/ᴜᴅᴀʀ′ᴀᴛ′ 'hit'

ASPECT	imperfective	perfective
INFINITIVE	udar′a-t′	udar′i-t′
GERUND	udar′a-ja	udar′i-v
IMPERATIVE	udar′a-j(te)!	udar′(te)!
TENSE		
present/future	udar′a-ju, -eš, …	udar′-u, -iš
L-PTCP	udar′a-l, -a, -o, -′i	udar′i-l, -a, -o, -′i
REFLEXIVE	udar′at′-s′a, etc	<none>

and the l-participle do not show any tense oppositions. The subject agreement shown by l-participles differs from that of finite verb forms in that it is defined in terms of MPSs proper to predicative adjectives, not those of finite verbs.

Table 3: CONTENT feature array

ASPECT	{ipfv,pfv}
VFORM	INFINITIVE
	TENSE:{prs, fut, pst}
	IMPERATIVE:{sg, pl}
	CONDITIONAL:{yes, no}
REFLEXIVE	{yes, no}
AGRSUBJ	PERSON:{1, 2, 3}
	NUMBER:{sg, pl}
	GENDER:{m, f, n}
VOICE	{ACTIVE, PASSIVE}
REPR	{⟨V,A⟩, ⟨V,Adv⟩}

In Tables 3, 4 I provide a summary list of the features which populate the CONTENT and FORM paradigms. I have provided only basic labels for the various MPSs. Ideally, we would want to know how they are grouped together, if at all, in the two paradigms. This is a difficult question, and I finesse it by just assuming what is effectively a list structure for the MPSs, with a number of dependency statements between them. Thus, I have not distinguished, say, a finite from a non-finite set of forms or constructions. However, for the purposes of giving a broad-brush characterization of the morphology~syntax mapping this is probably not a problem.

From this overview we can see that there is a very clear divide between the CONTENT paradigm MPSs required to describe the system as a whole and the FORM paradigm MPSs required to describe the individual word forms. These tables ignore the inflectional paradigms of the participles, of course, but adding them will just serve to emphasize the CONTENT~FORM disparity.

Table 4: FORM feature array

ASPECT	{ipfv,pfv}
VFORM	INFINITIVE
	TENSE:{prs-fut}
	IMPERATIVE:{sg, pl}
	L-PTCP
REFLEXIVE	{yes, no}
AGRSUBJ	PERSON:{1, 2, 3}
	NUMBER:{sg, pl}
	GENDER:{m, f, n}
REPR	{⟨V,A⟩, ⟨V,Adv⟩}

Russian has four participles, active~passive and perfective~imperfective, which are typical attributive modifiers with the agreement morphosyntax of standard adjectives. However, they retain a variety of verb properties, making them into typical examples of mixed categories. Thus, in (6a), the imperfective active participle *upravlʹajuščij* takes a temporal PP adjunct and assigns instrumental case to its complement, just like the finite verb (6b).

(6) a. (čelovek-a), upravlʹa-jušč-ego v tečenʹie mnogo let
(the.person[M]-GEN.SG) run-PRSPTCP-M.GEN.SG in course many of.years
mestnoj školoj
local.INSTR school.INSTR

'of (the person) (who has been) running the local school for many years'

b. Ivanov upravlʹaet v tečenʹie mnogo let mestnoj školoj
Ivanov runs in course many of.years local.INSTR school.INSTR

'Ivanov has been running the local primary school for many years'

The participles are often described as present/past tense forms, but their semantics is essentially aspectual and they fit somewhat better into the overall verb scheme if they are treated as perfective~imperfective pairs. They are summarized in Table 5 (cf Wade 1992: 361).

Table 5: Russian participles of the verb UPRAVʹITʹ/UPRAVLʹATʹ 'control'

Aspect	imperfective	perfective
Active	*upravlʹaju-šč-*	*upravʹi-vš-*
Passive	*upravlʹa-em-*	*upravlʹ-on(n)-*

Perfective aspect participles, for semantic reasons, usually only have past time reference.[8] The imperfective participles realize a time relative to the main verb of the clause, so that 'present tense' is a particularly misleading label for these forms (Wade 1992: 375).

(7) Ja vʹidel/vʹižu sobak-u,   bega-jušč-uju   po beregu
     I   saw/see   dog[F]-ACC.SG run-ACTPRSPTCP-F.ACC.SG along the shore

     'I saw/see the dog running along the shore'

The participle differs from the finite form in this respect. Example (8) would only be possible with either the meaning '(dog) which usually runs along the shore' or as a somewhat marked form of the historic present (cf *Russkaja Grammatika I*: 665).

(8) Ja vʹidel sobak-u,   kotor-aja   begaet po beregu
     I   saw   dog[F]-ACC.SG which-F.NOM.SG runs   along the shore

     'I saw the dog which is running along the shore'

In this respect, Russian participles are just like their English counterparts, of course.

The participles have a number of properties aligning them with verbs. In addition to realizing the purely verbal (eventive) categories of tense-aspect-voice, the active participles can take reflexive forms, either as reflexive variants of non-reflexives, inheriting all the semantics of the reflexive forms, *upravlʹatʹ(sʹa)* 'manage, control' ~ *upravlʹajuščijsʹa*, or as inherent reflexives (with no non-reflexive counterpart), *bojatʹsʹa 'fear'* ~ *bojaščijsʹa*. As we have seen, syntactically, the participles retain the verb's argument structure, including quirky case assignment to complements, such as instrumental in the case of *upravlʹajuščijsʹa* (see examples (6)) and genitive in the case of *bojaščijsʹa*.

On the other hand, the participles have a number of adjectival properties. The most salient morphosyntactic property is that of attributive adjective agreement together with the morphological property of belonging to a well-defined adjectival inflectional class. (As attributive modifiers the participles can be restrictive or non-restrictive, like other attributive modifiers, including relative clauses.) However, they can also be used as predicates with the copula BYTʹ 'be' or with "semi-copulas" such as STATʹ 'become', KAZATʹSʹA 'seem', OSTATʹSʹA 'remain', and others. Most commonly it is the perfective passive participles that can be used as predicates but active participles can also be found in this role (Švedova 1980b: 291, §2346). The passive (though not the active) participles can also appear as predicates, in the so-called short form, a typically adjectival property.[9]

(9) užin   uže   poda-n
     supper.M.SG already serve-PASSPTCP.M.SG

     'Supper has already been served'

---

[8] But see the counter-examples in the Academy of Sciences grammar *Russkaja Grammatika, I*: 667, *predʹjavlʹajuščij* 'presenting', *vzvolnujuščij* 'exciting', *sdelajuščij* 'doing', *smoguščij* 'being able'.

[9] Present active participles can be used in the short form, however, when they are converted into true, qualitative, adjectives (Švedova 1980a: 666).

Given this brief descriptive summary of the basic facts of Russian participles we can turn to the central questions: how do we represent participles in a formal, constraints-based grammar with an inferential-realizational morphology? How do these representations relate to the inflection~derivation divide and the issue of split morphology?

5 Paradigm linkage rules for Russian

In the extension to GPFM presented here, the FORM paradigm and the CONTENT paradigm are modelled by the attributes f-MORSIG, c-MORSIG. However, those types of lexical relatedness which modify the MPSs of a representation, such as transpositions, will ipso facto modify the content of Stump's FORM/ CONTENT paradigms and the f-/c-MORSIG attributes. The GPF for such types of relatedness therefore has to specify that novel content, by stipulation, if necessary. In the rules I propose below I show how this can be achieved for Russian conjugation. The reference to MORSIG is taken to mean c-MORSIG by default and by default, this is identical to f-MORSIG.

I shall begin by specifying the CONTENT and FORM MORSIG attribute for non-participial verb categories. We need to define the f-MORSIG in part in terms of the c-MORSIG and in part independently. The default is the identity mapping from c-MORSIG to f-MORSIG. The c-MORSIG list shown earlier in Table 1 is defined for any lexical representation whose ARG-STR includes the E semantic function role. The MPSs that are shared across the CONTENT and FORM paradigms of Russian verbs are fairly limited (cf. Tables 1 and 2 above). They are (ignoring for the present the participles and gerunds): ASPECT:{ipfv, pfv}, VFORM:{INFINITIVE, IMPERATIVE:{sg, pl}}, VOICE:{act, pass}, and AGRSBJ:{PERSON/NUMBER/GENDER}.

The c-features ASPECT, VFORM:{INFINITIVE, IMPERATIVE:{sg, pl}} have relatively straightforward f-feature correspondents. I shall ignore INFINITIVE and IMPERATIVE for present purposes. AGRSUBJ is also a FORM property but with some complications and I return to it when I discuss the l-participle.

The status of the TENSE feature is a little unclear. At the CONTENT level there are clearly three values, *prs, fut, pst*, but only the *prs* and *fut* values have a FORM correspondent, and even then the value of the FORM correspondent is a composite *prs/fut*, and therefore not a direct correspondent of either c-[TENSE:prs] or c-[TENSE:fut]. The c-TENSE:{pst} property is realized by the morphomic l-participle, and not by any dedicated f-TENSE:{pst} property. I shall therefore assume a univalent FORM property, TNS, itself a value of VFORM, realizing c-TENSE:{prs, fut} depending on the value of ASPECT. This replaces the atom-valued TENSE:{present-future} shown in Table 2 above. The property VOICE:{act, pass} is intriguing. It is a CONTENT property of the verb system but it is not a FORM property of any part of the verb system proper, outside of the participle subsystem. However, the participles distinguish active and passive sets, and the perfective passive participle is actually part of the periphrastic exponence of the syntactic (CONTENT) VOICE property. Moreover, the two passive participles have the passive SYN|ARG-STR representation, namely, …⟨E⟨(x), y, …⟩, where (x) denotes the demoted active subject argument role. Therefore, VOICE is both part of the CONTENT and the

FORM paradigm, albeit with a somewhat complex realization, which will require the f-MORSIG attribute to be modified by a feature co-occurrence statement restricting f-VOICE to the participles. The other CONTENT paradigm features are represented by forms which are effectively periphrastic. It is for these reasons that the FORM MPSs have to be defined independently, as shown below.

One attribute that is shared across FORM/CONTENT paradigms is REPRESENTA-TION. The only reason for a language to define a true transpositional category is to allow a lexeme to assume the syntactic distribution of a word of a different class, so REPR clearly has to be a CONTENT feature. In languages such as Russian, in which participles are marked morphologically, this also means that REPR is a FORM feature. I shall assume that the REPR attribute as applied to verbs has three values. The first is *plain* (or $\langle V,V \rangle$), the 'identity representation', and the default value. Where no indication of REPR is given the default is to be understood. The second value of REPR is $\langle V, A \rangle$, defining the four participles. The third value is $\langle V,Adv \rangle$,[10] defining the imperfective and perfective gerunds. I shall ignore the gerunds for simplicity of exposition. This means that I shall only mark the participles explicitly.

We now define the content of the c-MORSIG attribute by reference to the verb's SYN value. The aspects, and voice are defined with the mapping shown informally in (10). This will apply to any representation whose ARG-STR contains the E sf role. In practice, this means verbs, participles, and gerunds.

(10)   ... E ... $\Rightarrow$ ASPECT, VOICE $\subset$ MORSIG

I return to the c-MORSIG|CONDITIONAL, TENSE:{pst} mappings below when I discuss the status of the l-participle.

The feature array defining participles is given in Fig. 1. This is the "derived" MORSIG attribute for any lexical representation defined by the feature REPR:$\langle V,A \rangle$. The property sets labelled $\boxed{1}$ in Fig. 1 come from the MORSIG attribute of the base verb by virtue of (10). The property sets $\boxed{3}$, $\boxed{4}$ are those which ensure that the participles are inflected like adjectives. I return to these when I have introduced adjective inflection.

The key to my analysis of transpositions is to incorporate part of the analysis of derivational morphology into the definition of the transposition's entry. Given a feature pairing $\langle £,\rho \rangle$, where £ is a lexemic index and ρ contains a value of REPR (for instance, [REPR:$\langle V,A \rangle$] for participles), the GPF applied to this pairing leaves the LI and the SEM representations of the base unchanged, but enriches the SYN|ARG-STR attribute by creating a complex semantic function role $\langle A_i \langle E \langle x_i, ... \rangle \rangle \rangle$. The coindexation guarantees that the noun head modified by the participle is identified with the highest thematic argument of the base verb's argument array, i.e. its SUBJECT. The FORM function component of the GPF defines the stem set for the participle, but underspecifies all other FORM information, including the MORSIG (and the CONTENT paradigm MORSIG is also underspecified by the SYN function).

---

[10] The semantic function role label Adv stands proxy for whatever the appropriate way is of defining adverbs as distinct from adjectives.
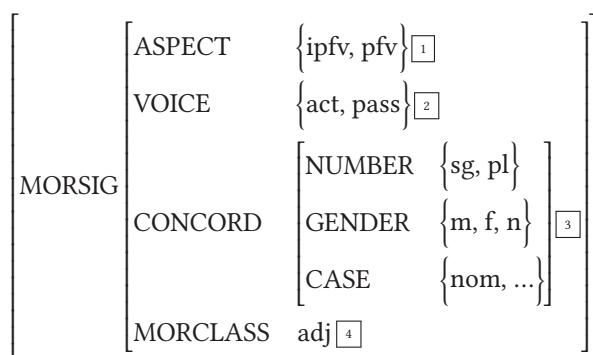
$$\begin{bmatrix} \text{MORSIG} & \begin{bmatrix} \text{ASPECT} & \{\text{ipfv, pfv}\}_{\boxed{1}} \\ \text{VOICE} & \{\text{act, pass}\}_{\boxed{2}} \\ \text{CONCORD} & \begin{bmatrix} \text{NUMBER} & \{\text{sg, pl}\} \\ \text{GENDER} & \{\text{m, f, n}\} \\ \text{CASE} & \{\text{nom, ...}\} \end{bmatrix}_{\boxed{3}} \\ \text{MORCLASS} & \text{adj}_{\boxed{4}} \end{bmatrix} \end{bmatrix}$$

Figure 1: Feature structure for Russian participles

The lexemic entry for a typical transitive Russian verb such as UDAR´IT´/UDAR´AT´ is that shown in Fig. 2, where $\mathcal{V}$ stands for the verb's lexemic index. The lexemic entry's value for LI is just the LI of the lexeme, of course (the GPF here does not describe a process of derivational morphology). Note that for this lexeme the SYN attribute, too, is completely underspecified, lacking even the ARG-STR attribute. The value of that attribute is determined by default from the *Event* ontological type of the SEM representation. However, the REPR feature which defines transpositions introduces a realization rule which has to be defined over a specified ARG-STR representation. Therefore, the ARG-STR attribute has to be part of the lexeme's (SYN attribute's) morpholexical signature (inflection-like argument structure alternations such as passive or antipassive impose a similar requirement). In this respect, the transpositions are like inflection and not like derivation.

We can informally state the realization rule which defines adjectival representations of lexemes (transpositions-to-adjective) as a function α from ARG-STR representations to ARG-STR representations, as shown in (11), where (11a) defines a participle's ARG-STR and (11b) defines that of a relational adjective.

(11) a. $\alpha(\langle E\langle x,...\rangle\rangle) = \langle A_i\langle E\langle x_i,...\rangle\rangle\rangle$

  b. $\alpha(\langle R\rangle) = \langle A \langle R\rangle\rangle$

Given the lexemic entry in Fig. 2 and the realization rules for Russian morphology, the GPF for the imperfective active participle, *udar´ajušč-* will map to a partially underspecified lexical representation, as shown in Fig. 3. A representation such as this is the 'quasi-lexeme' discussed earlier. Like a simple lexemic entry, or an entry defined by a derivational GPF, it needs to have its MORSIG attributes specified in order to be inflectable. I turn now to how those MORSIG entries are defined.

The partially specified MORSIG we need to be able to define for participles is that shown in Fig. 4. The ASPECT/VOICE properties are shared with verbs. This can be achieved by writing the rules defining the MORSIG of verbs and participles in such as way as to refer either to the "outermost" E semantic function role of the ARG-STR at-

$$\mathrm{f}_{form}\left(\langle\mathcal{V},\,u\,\rangle\right)\quad=\quad \begin{bmatrix}\mathrm{STEM0}\begin{bmatrix}\mathrm{PHON} & udar' \\ \mathrm{MORCLASS} & \mathrm{V\,|\,CONJ2}\end{bmatrix} \\ \mathrm{STEM1}\begin{bmatrix}\mathrm{PHON} & udar'aj \\ \mathrm{MORCLASS} & \mathrm{V\,|\,CONJ1}\end{bmatrix}\end{bmatrix}$$

$$\mathrm{f}_{syn}(\langle\mathcal{V},\,u\,\rangle) \quad = \quad u$$

$$\mathrm{f}_{sem}(\langle\mathcal{V},\,u\,\rangle) \quad = \quad \lambda x,y[\mathrm{HIT}(x,y)]$$

$$\mathrm{f}_{li}(\langle\mathcal{V},\,u\,\rangle) \quad = \quad u$$

Figure 2: Lexemic entry for UDAR'IT'/UDAR'AT'

$$\mathrm{FORM}(\langle\mathcal{V},\,\{\mathrm{REPR}{:}\langle\mathrm{V,A}\rangle\}\rangle) \quad = \quad \begin{bmatrix}\mathrm{STEM0\,|\,PHON} & udar'ajušč \\ \mathrm{MORSIG} & u\end{bmatrix}$$

$$\mathrm{SYN}(\langle\mathcal{V},\,\{\mathrm{REPR}{:}\langle\mathrm{V,A}\rangle\}\rangle) \quad = \quad \begin{bmatrix}\mathrm{ARG\text{-}STR} & \langle\mathrm{A}_i\langle\mathrm{E}\langle x_i,y\rangle\rangle\rangle \\ \mathrm{MORSIG} & u\end{bmatrix}$$

$$\mathrm{SEM}(\langle\mathcal{V},\,\{\mathrm{REPR}{:}\langle\mathrm{V,A}\rangle\}\rangle) \quad = \quad \text{identity function}$$

$$\mathrm{LI}(\langle\mathcal{V},\,\{\mathrm{REPR}{:}\langle\mathrm{V,A}\rangle\}\rangle) \quad = \quad \text{identity function}$$

Figure 3: Underspecified entry for *udar'ajušč(ij)*

$$\begin{bmatrix}\mathrm{MORSIG}\begin{bmatrix}\mathrm{ASPECT:\,\{u\}} \\ \mathrm{VOICE:\,\{u\}} \\ \mathrm{CONCORD:\{u\}} \\ \mathrm{MORCLASS:\,ADJ\,|\,DECL1/2}\end{bmatrix}\end{bmatrix}$$

Figure 4: MORSIG for Russian participles

tribute, or the "embedded" E role found with participles.[11] The CONCORD attribute, $\boxed{3}$, comes from the generic c-MORSIG of an adjective, shown in (12).[12]

(12)   ... A ... $\Rightarrow$ [CONCORD:{NUMBER, GENDER, CASE}] $\subset$ MORSIG

The [MORCLASS adj] specification, $\boxed{4}$, is strictly speaking a stipulation, except that in Russian (in contrast to, say, Latin) all participles belong to the default adjectival class, DECL1/2.

Given this machinery we can now account for the transpositional mixed category of participle by application of the (quasi-inflectional) generalized paradigm function applying to a verb and delivering its participial forms, triggered by the [REPR] feature. These representations will be underspecified for (adjectival) CONCORD features. For concreteness, consider the imperfective active participle, *udar′ajušč-*. Given $\rho = \{$REPR:$\langle$V,A$\rangle$, ASPECT:ipfv $\boxed{1}$, VOICE:act $\boxed{2}\}$, then for $\mathcal{V} = $ UDAR′IT′/UDAR′AT′, the GPF($\langle\mathcal{V},\rho\rangle$) will apply to a lexical representation which is derived from the lexemic entry for UDAR′IT′/UDAR′AT′ whose MORSIG attributes have been fully specified, allowing the lexeme to be inflected (in the broad sense of this term, including "inflection" for participle formation). This GPF will deliver a partially underspecified lexical representation for the participle. The GPF will specify the participle's stem form(s), the ASPECT, VOICE features which define that particular participle, and the enriched ARG-STR attribute with complex semantic function role.

The lexical representation which is input to the GPF is shown in Fig. 5 and the lexical representation of the participle is shown in (13). In Fig. 5, STEM0 denotes the perfective stem, which is effectively the lexeme's root. CONJ2 is second conjugation, and this means that most inflectional rules will be defined over another stem, *udar′i-*, derived by regular rules of stem formation. STEM1 denotes the imperfective stem, a member of CONJ1, the first conjugation, whose inflectional stem is therefore *udar′aj-*. Attributes which belong to both FORM and SYN (CONTENT) paradigms are tagged to make them more easily identifiable.

(13) $f_{form} = $ STEM-iap $\oplus$ šč = udar′aju-šč-
        $f_{syn} = $ ARG-STR:$\langle A_i \langle E \langle x_i, y \rangle \rangle \rangle$

where STEM-iap denotes the imperfective active participle stem, derived from STEM1.

We have now achieved our goal of defining participles. In effect, we have defined the lexemic entry for a class of adjectives, whose peculiarity is that they are marked for verbal voice and aspect and they share their semantics and lexemic index with a parent verb.

It is instructive to compare the behaviour of the true participles with that of the l-participle. While the true participles are mixed categories, the l-participle is essentially a verb form with unusual agreement morphology. It is not entirely clear how best to

---

[11] Russian action nominals are transpositional lexemes and not true transpositions.

[12] Members of the semantically defined class of quality or scalar adjectives will also have the feature COMPARISON added to their MORSIG to define comparative/superlative morphology.

$$
\begin{bmatrix}
\text{FORM} & \begin{bmatrix}
\text{PHON} & \begin{bmatrix}
\begin{bmatrix} \text{STEM0 PHON} \mid \text{udar}' \\ \text{MORCLASS V} \mid \text{CONJ2} \end{bmatrix} \\
\begin{bmatrix} \text{STEM1} \\ \text{MORCLASS V} \mid \text{CONJ1} \end{bmatrix}
\end{bmatrix} \\
\text{MORSIG} & \begin{bmatrix}
\text{ASPECT:}\textit{ipfv, pfv}\ \boxed{1} \\
\text{VOICE:}\textit{act, pass}\ \boxed{2} \\
\text{VFORM:}\begin{bmatrix} \text{TNS} \\ \text{LPTCP} \\ \ldots \end{bmatrix} \\
\text{CONCORD:}\{\text{NUMBER, PERSON, GENDER}\}\ \boxed{3} \\
\text{REPR:}\{\langle V,A \rangle, \langle V,Adv \rangle\}\ \boxed{4} \\
\text{Other verb FORM MPSs}\ \boxed{6}
\end{bmatrix}
\end{bmatrix} \\
\text{SYN} & \begin{bmatrix}
\text{ARG-STR:} & \langle E \langle x_i, y \rangle \rangle\ \boxed{5} \\
\text{MORSIG} & \begin{bmatrix}
\text{ASPECT:}\{\textit{ipfv, pfv}\}\ \boxed{1} \\
\text{VOICE:}\{\textit{act, pass}\}\ \boxed{2} \\
\text{TENSE:}\{\textit{pst, prs, fut}\} \\
\text{CONCORD:}\{\text{NUMBER, PERSON, GENDER}\}\ \boxed{3} \\
\text{REPR:}\{\langle V,A \rangle, \langle V,Adv \rangle\}\ \boxed{4} \\
\text{ARG-STR:}\langle \text{sf role } \langle \theta \text{ array} \rangle \rangle\ \boxed{5} \\
\text{Other verb CONTENT MPSs}\ \boxed{6}
\end{bmatrix} \\
\ldots
\end{bmatrix} \\
\text{SEM} & \lambda x,y[{}_{Event}\text{HIT}(x,y)] \\
\text{LI} & \textsc{udar}'\textsc{it}'/\textsc{udar}'\textsc{at}'
\end{bmatrix}
$$

Figure 5: Lexical representation of the lexeme UDAR'IT'/UDAR'AT' 'to hit'

account for the peculiar agreement properties of the l-participle in the past tense and conditional constructions, but the simplest (if somewhat crude) way to do this is to assume that verbs in general agree with their subjects in person, number, gender features, but that in the past/conditional forms agreement is restricted to number, gender, while in the present tense forms it is restricted to person, number.[13]

I conclude, then, that the l-participle is a verb form that inflects just like a (subtype of) predicative (short-form) adjective. Assuming a feature ADJDECL covering adjectival morphology generally, we can distinguish several sub-types of declension, including that for the predicative adjectives, [ADJDECL:predadj]. The l-participles will belong to a subtype of this class, [ADJDECL:predadj:lptcpdecl]. The ADJDECL feature is part of the MORSIG attribute of ordinary adjectives, as defined by the paradigm linkage rule in (14).[14]

(14)　… A … $\Rightarrow$ ADJDECL $\subset$ MORSIG

We also compare true participles with qualitative adjectives derived by conversion from participles. These are a type of transpositional lexeme. The theoretical significance of this type of lexical relatedness for current morphological models was first identified, as far as I am aware in Spencer (2013: 275), where I discuss English words such as PREPOSITIONAL, from PREPOSITION. These look like relational adjectives (noun-to-adjective transpositions), because their lexical semantic content seems to be identical to that of their base noun (*prepositional phrase* means the same as the compound *preposition phrase*, for instance). However, the English adjectives are syntactically opaque: *monosyllabic prepositional phrase* does not mean 'phrase headed by a monosyllabic preposition' but only 'monosyllabic phrase headed by a preposition' (though that interpretation is possible for the compound, *monosyllabic preposition phrase*). In other words, the adjectival expression only has the structure *monosyllabic [prepositional phrase]*, not *[monosyllabic preposition]al phrase*.

English has a large number of qualitative adjectives derived by conversion from participles, which are also instances of transpositional lexemes (Spencer 2016a): *amazed/amazing, bored/boring, challenged/challenging, interested/interesting, ….* In very many cases it is not possible to identify a meaning difference between the adjective and the etymological verb base: *This book bores me/This book is boring*. Russian participles likewise are often converted into qualitative adjectives: *potrˊasajuščij* 'amazing', *vyzyvajuščij* 'provocative, defiant, challenging'. In Russian it is often possible to determine that a word with the shape of a participle is actually an independent adjective, since only true adjectives have the short predicative form: *uspexi potrˊasajušči* (plural, from *potrˊasajušč*)

---

[13] A more sophisticated approach might define agreement in a morphology-driven fashion by stating that the agreement features that the syntax can manipulate are restricted to those that can be expressed by a particular morphological form, so that it is the morphological MORSIG that determines which features trigger agreement, even in identical syntactic positions.

[14] As adjectives, we might expect the participles to have predicative forms, too. This is true, however, only for the passive participles, especially the perfective passive, which has a special stem form ending in a singleton /n/, *napˊisan* 'written', in contrast to the attributive form with geminated /nn/: *(v speške) napˊisannaja (zapˊiska)* '(a hastily) written (note)'.

'the-progress is-amazing' (lit. 'the-successes are-amazing'), *ego poveden´ie vyzyvajušče* 'his behaviour is-defiant'. True participles do not have the predicative form (Russkaja Grammatika II, p.666). English converted participles fail to inherit the complementation properties of the base verb: *The obstacle course challenged the stamina of the athletes∼ The obstacle course was very challenging (*the stamina of the athletes).* Russian transpositional adjective lexemes behave likewise. The transitive base verb POTR´ASAT´ 'to amaze' gives us *uspexi potr´asal´i nas* 'the-successes amazed us', but the (active) transpositional lexeme cannot take a direct object: **uspexi potr´asajušči nas.* The true participle remains transitive: *potr´asajuščie nas uspexi* 'progress (successes) which amaze us'.

We thus have a double dissociation of properties: on the one hand, we have the l-participle which has the form of a predicative adjective but which realizes finite (tense/mood) verb properties and retains the full complementation properties of the verb, and on the other hand we have participle-like adjectives which, while allowing predicative adjective forms, lack all the crucial morphosyntax of verbs. In between we have the true participles, with the external morphosyntax of an adjective but the complementation properties of the base verb.

6 Conclusions: Transpositions and split morphology

I have argued in this paper for a view of morphosyntax which recognizes a word/phrase distinction and, given that, a distinction between (abstract) lexemes and (concrete) inflected word forms of those lexemes, valid for very nearly all known languages. I have also assumed that languages can increase their stock of lexemes by means of derivational morphology, and that in some cases this is sufficiently regular to be regarded as paradigmatic, hence, part of the grammatical system proper. The inflection/derivation distinction is controversial, however, because it is often difficult to know where the boundary actually lies and where to place intermediate types of lexical relatedness. The transpositions, as exemplified by the Russian participial system discussed here, represent a particularly troublesome case-in-point.

Participles and other transpositions are often treated as a type of derivational morphology, because they involve a shift in word class, but this is a wrong characterization. Participles are part of a verb's paradigm, they are not a type of lexical stock expansion (Beard 1995). Nonetheless, participles inflect like adjectives, not like verbs and thus seem to straddle the inflection/derivation divide in a way that calls that very distinction into question, and, on the face of it, even provides support for models in which all morphology is just syntax by other means (Minimalism/Distributed Morphology) or in which all syntax is just morphology by other means (American morphemics). Here I have claimed that, on the contrary, we can only make sense of participles against a set of background assumptions that contradict monolithic models in which there is no autonomous morphology module (Aronoff's "morphology-by-itself"), so that morphology is no more than syntax by other means. The crucial observation is that derivational morphology induces a type of lexical opacity which is lacking in transpositions, which, by contrast, show the kind of lexical transparency associated with inflected forms.

In the GPFM model, canonical derivation is a relation between maximally underspecified (minimally redundant) lexical representations (lexemic entries), consisting of a specification of the basic root form (FORM|PHON) and the ontological/semantic representation (SEM). The morphosyntactic properties are then projected from these by default mappings. Canonically, derivation enriches the PHON and the SEM representations, and the Default Cascade then specifies the morphosyntactic properties of the derived lexeme. One consequence is that there will then be no "trace" left of the morphosyntactic properties of the base lexeme, such as its word class or its argument structure. This automatically guarantees most of the lexical opacity/lexical integrity effects familiar from the literature. In some (noncanonical) cases it may be necessary to stipluate overrides of lexical information as part of the derivational Generalized Paradigm Function. This might be true if, for instance, a base lexeme belongs to a lexical category which is not the default for its ontological class. For instance, a language with a distinct lexical category of (qualitative) adjective may also have non-derived stative verbs whose denotations are of the ontological type Property, and which by the Default Cascade should be adjectives, not verbs. Such a verb would have to have its ARG-STR|⟨E …⟩ value prespecified. If such a lexeme were the input to a verb-to-noun derivational function (nominalization), then that ARG-STR would have to be overridden by the nominalization function, replacing it wholesale with the ARG-STR|⟨R⟩ value. Exactly how such cases are to be handled has to remain a matter for future research.

No such opacity is found with canonical inflection: in general, the syntax treats a noun as a noun no matter what its number, case, definiteness, … value. Thus, although a locative case marked noun would normally be restricted to functioning as an adjunct or the complement of a class of adpositions, it can still be modified by an adjective, just like any other noun form, so that for a noun 'house', *new house-LOC* means 'at a new house'. In this respect, a locative case marked noun typically differs from a derived denominal lexeme denoting a location. Many languages have a denominal derivational marker meaning 'place where there is/are NOUN, place associated with NOUN': N-PLACE. Typically, the base noun, N, is not accessible to morphosyntax, so that an expression such as *new house-PLACE* could only mean 'new place where there is a house/are houses' and not 'place where there is/are a new house(s)'. Thus, inflection differs from derivation in being lexically transparent.

The significance of these rather obvious points about inflected forms is that we are far less able to make similarly categorical claims where transpositions are concerned. A participle behaves in the syntax *to some extent* as though it were an inflected verb form, but not entirely. In GPFM the lexical representation of a transposition exhibits transparency by virtue of being a member of the base lexeme's (extended) paradigm, that is, by bearing the same LI as the base. In this respect it is not an autonomous lexeme. On the other hand, the transposition exhibits the external syntax of a distinct (derived) lexical category. The extension to the GPF proposed here permits us to model this "inflectional-paradigm-within-a-paradigm" effect in a way that reflects the lexical transparency of the transposition while also allowing us to state restrictions on full transparency as a restriction on the MORSIG of the transposition.

A crucial aspect of the analysis is the distinction between FORM/CONTENT properties (m-/s-features). Without at least this level of differentiation we cannot make sense of Russian participles and we cannot distinguish them from the verbal l-participle form or the transpositional lexemes. A second crucial aspect of the analysis of any type of transposition is the Lexemic Index (LI). In the general case, there is no combination of lexical properties that will uniquely serve to individuate lexemes, but it is nonetheless essential to impose such an individuation to account for the patterns of lexical relatedness that are found across languages, specifically, to distinguish true participles from departicipial converted adjectives (transpositional lexemes).

The combination of the FORM/CONTENT paradigm distinction (or its equivalent) and the Lexemic Index allow us to define not just canonical inflection but also non-canonical intra-lexemic types of relatedness such as that shown by transpositions. That combination also serves to reconstruct the split in the morphology argued for originally by Anderson. Indeed, split morphology is entailed by this set of assumptions, except that in a constraints-based model the split is defined in terms of access: inflectional morphology is that which permits syntax to retain access to the properties of the base lexeme (lexical transparency), while derivational morphology permits no such access (lexical opacity/integrity).

The more articulated view of morphosyntax proposed here allows us to pose a question which was not at the forefront of debate over the question of split morphology, as far as I am aware: which side of the split do transpositions fall on? The answer, given the foregoing, is "both". The transposition is inflectional by virtue of preserving the base's LI and by virtue of the, at least partial, transparency of the base's properties. It is derivational by virtue of the fact that it defines the paradigm of a quasi-lexeme, within the paradigm of the base. But it would be difficult to make sense of this conclusion without assuming the basic split in the first place.

Appendix: Illustration of the lexical representations assumed

To specify, say, the 3sg form *drives*, the GPF(⟨DRIVE,{3sg}⟩) applies to the output of Figure 6 to specify the FORM value |STEM0⊕z|, leaving other aspects of the representation unchanged. This is equivalent to the operation of the paradigm function in PFM1 and the output of the ***Corr*** function in PFM2.

$$
\begin{bmatrix}
\text{FORM} & \text{STEM0} \;\; |\,\text{draɪv}\,| \\
\text{SYN} & - \\
\text{SEM} & [_{Event}\; \lambda x, y.\mathbf{drive}(x,y)] \\
\text{LI} & \textsc{drive}
\end{bmatrix}
\quad \Rightarrow
$$

$$
\begin{bmatrix}
\text{FORM} &
\begin{bmatrix}
\text{STEM0} & |\,\text{draɪv}\,| \\
\text{MORSIG} &
\begin{bmatrix}
\text{TNS} & \{\text{prs, pst}\} \\
\text{VFORM} & \{\text{ing-form, ed-form, base}\} \\
\text{AGR} & \{\text{3sg, none}\}
\end{bmatrix}
\end{bmatrix} \\[1em]
\text{SYN} &
\begin{bmatrix}
\text{SYNCAT} & \text{V} \\
\text{ARG-ST} & \langle E\langle x,y\rangle\rangle \\
\text{MORSIG} &
\begin{bmatrix}
\text{TNS} & \{\text{prs, pst}\} \\
\text{ASP} & \{\text{simple, prog, perf}\} \\
\text{AGR} & \{\text{3sg, none}\} \\
\dots
\end{bmatrix}
\end{bmatrix} \\[1em]
\text{SEM} & [_{Event}\; \lambda x, y.\mathbf{drive}(x,y)] \\
\text{LI} & \textsc{drive}
\end{bmatrix}
$$

Figure 6: Illustration of the application of the GPF for standard inflection.

(15) DRIVE ⇒ DRIVER

Where δ = SubjNom, Σ = [$_{Event}$ λx,y.**drive**(x,y)], GPF(⟨DRIVE,δ⟩) =

$$
\begin{bmatrix}
\text{FORM} & \text{STEM0} \ | \text{draɪv} | \\
\text{SYN} & — \\
\text{SEM} & Σ \\
\text{LI} & \text{DRIVE}
\end{bmatrix}
\Rightarrow
$$

$$
\begin{bmatrix}
\text{FORM} & [\text{STEM0} & \text{STEM0(DRIVE)} \oplus | ə | \\
\text{SYN} & — \\
\text{SEM} & [_{Thing} \ λx.\text{PERSON}(x) \wedge Σ \\
\text{LI} & δ(\text{DRIVE})
\end{bmatrix}
$$

Figure 7: Illustration of the application of the GPF for standard derivation

The output of this GPF then undergoes specification of MORSIG by the Default Cascade, which defines the derived lexeme as a syntactic and morphological noun.

Abbreviations

AGRSUBJ	subject agreement	MORCAT	morpholexical category
act	active		
ARG-STR	argument structure	MORCLASS	morphological class
		MORSIG	morpholexical signature
CONJ	conjugation		
DECL	declension	MPS	morphosyntactic property set
GPF	Generalized Paradigm Function		
		pass	passive
		PF	Paradigm Function
GPFM	Generalized Paradigm Function Morphology	PFM	Paradigm Function Morphology
		pfv	perfective
		PHON	phonological form
ipfv	imperfective	REPR	REPRESENTATION
LI	lexemic index	SEM	semantics
L-PTCP, l-ptcp	l-participle	SYN	syntax

Acknowledgements

References

Ackerman, Farrell & Phil LeSourd. 1997. Toward a lexical representation of phrasal predicates. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates*, 67–106. Stanford, CA: CSLI Publications.

Anderson, Stephen R. 1977. On the formal description of inflection. In Woodford A. Beach, Samuel E. Fox & Shulamith Philosoph (eds.), *Papers from the thirteenth regional meeting of the Chicago Linguistic Society*, 15–44. Chicago: Chicago Linguistic Society.

Anderson, Stephen R. 1982. Where's Morphology? *Linguistic Inquiry* 13(4). 571–612.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Anderson, Stephen R. 2015. The morpheme: Its nature and use. In Matthew Baerman (ed.), *The Oxford handbook of inflection*, 1–34. Oxford: Oxford University Press.

Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT Press.

Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2005. *The syntax-morphology interface: A study of syncretism*. Cambridge: Cambridge University Press.

Baerman, Matthew, Greville G. Corbett, Dunstan Brown & Andrew Hippisley (eds.). 2007. *Deponency and morphological mismatches*. Oxford: Oxford University Press.

Beard, Robert. 1995. *Lexeme-morpheme base morphology*. Stony Brook, NY: SUNY Press.

Bonami, Olivier. 2015. Periphrasis as collocation. *Morphology* 25. 63–110.

Bonami, Olivier & Gregory Stump. 2016. Paradigm Function Morphology. In Andrew Hippisley & Gregory T. Stump (eds.), *The Cambridge handbook of morphology*, 449–481. Cambridge: Cambridge University Press.

Booij, Geert. 1994. Against split morphology. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1993*, 27–49. Dordrecht: Kluwer Academic Publishers.

Brown, Dunstan P., Marina Chumakina, Greville G. Corbett, Gergana D. Popova & Andrew Spencer. 2012. Defining 'periphrasis': Key notions. *Morphology* 22(2). 233–275.

Gerritsen, Nellike. 1990. *Russian reflexive verbs: In search of unity in diversity*. Amsterdam: Rodopi.

Haspelmath, Martin. 1996. Word-class changing inflection and morphological theory. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1995*, 43–66. Dordrecht: Kluwer Academic Publishers.

Jackendoff, Ray S. 1990. *Semantic structures*. Cambridge, MA: MIT Press.

Lieber, Rochelle. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press.

Lowe, John. 2015. *Participles in Rigvedic Sanskrit: The syntax and semantics of adjectival verb forms*. Oxford: Oxford University Press.

Malchukov, Andrej L. 2004. *Nominalization/verbalization: Constraining a typology of transcategorial operations*. Munich: Lincom Europa.

Matthews, Peter H. 1972. *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation*. Cambridge: Cambridge University Press.

Miller, Philip H. & Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15(3). 573–639.

Nikolaeva, Irina A. 2008. Between nouns and adjectives: A constructional view. *Lingua* 118(7). 969–996.

Popova, Gergana & Andrew Spencer. Forthcoming. Stacked periphrases. In *Proceedings of the 11th Formal Description of Slavic Languages conference*.

Sadler, Louisa & Rachel Nordlinger. 2006. Case stacking in Realizational Morphology. *Linguistics* 44(3). 459–488.

Sadler, Louisa & Andrew Spencer. 2001. Syntax as an exponent of morphological features. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2000*, 71–96. Dordrecht: Kluwer Academic Publishers.

Sims, Andrea. 2015. *Inflectional defectiveness*. Cambridge: Cambridge University Press.

Spencer, Andrew. 1999. Transpositions and argument structure. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 1998*, 73–102. Dordrecht: Kluwer Academic Publishers.

Spencer, Andrew. 2002. Gender as an inflectional category. *Journal of Linguistics* 38(2). 279–312.

Spencer, Andrew. 2013. *Lexical relatedness: A paradigm-based model*. Oxford: Oxford University Press.

Spencer, Andrew. 2016a. How are words related? In Daniel Siddiqi & Heidi Harley (eds.), *Morphological metatheory*, 1–26. Amsterdam: John Benjamins.

Spencer, Andrew. 2016b. Individuating lexemes. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of LFG15*, 357–377. Stanford: CSLI Publications.

Spencer, Andrew. 2016c. Participial relatives in LFG. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of LFG15*, 378–398. Stanford: CSLI Publications.

Spencer, Andrew & Gregory Stump. 2013. Hungarian pronominal case and the dichotomy of content and form in inflectional morphology. *Natural Language and Linguistic Theory* 31. 1207–1248.

Stewart, Thomas & Gregory Stump. 2007. Paradigm Function Morphology and the morphology-syntax interface. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 383–421. Oxford: Oxford University Press.

Stump, Gregory. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.

Stump, Gregory. 2002. Morphological and syntactic paradigms: Arguments for a theory of paradigm linkage. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2001*, 147–180. Dordrecht: Kluwer Academic.

Stump, Gregory. 2006. A theory of heteroclite inflectional paradigms. *Language* 82. 279–322.

Stump, Gregory. 2016a. *Inflectional paradigms: Content and form at the syntax-morphology interface*. Cambridge: Cambridge University Press.

Stump, Gregory. 2016b. Paradigms at the interface of inflectional morphology. In Daniel Siddiqi & Heidi Harley (eds.), *Morphological metatheory*, 27–58. Amsterdam: John Benjamins.

Švedova, Natalija Ju. (ed.). 1980a. *Russkaja grammatika*. Vol. 1. Issued by the Academy of Sciences of USSR. Moskva: Izdatel'stvo nauka.

Thornton, Anna. 2012. Reduction and maintenance of overabundance. A case study on Italian verb paradigms. *Word Structure* 5. 183–207.

Wade, Terence. 1992. *A comprehensive Russian grammar*. Oxford: Blackwell Publishing.

Wechsler, Stephen. 2014. *Word meaning and syntax: Approaches to the interface*. Oxford: Oxford University Press.

Švedova, Natalija Ju. (ed.). 1980b. *Russkaja grammatika*. Vol. 2. Issued by the Academy of Sciences of USSR. Moskva: Izdatel'stvo nauka.