

5 *Formal framework and case studies*

As we saw in the previous chapter, the challenge for a formal analysis of syncretism is twofold: to represent the cross-linguistically more general patterns of syncretism in terms of feature structure, and to account for patterns which occur independently of feature structure. In this chapter we lay out a formal framework and demonstrate it with three cases studies. When introducing the formal framework, we show in §5.1 that inferential-realizational frameworks, such as Network Morphology and Paradigm Function Morphology, are capable of capturing syncretic patterns which may arise as the result of underspecification, or as the result of referrals (i.e. independently of feature structure). One advantage of such approaches is that referrals and underspecification can be used simultaneously. We shall see when we come to the second case study that this is just what is required for the avoidance morphology of Dalabon. Indeed, we show that generalized referrals – where sets of paradigm cells can refer to sets of paradigm cells – which frameworks such as Network Morphology allow for, are an important requirement.

We consider the relationship between underspecification (a feature structure-based constraint) and semantic naturalness. As we saw in Chapter 4: §4.2.1, underspecification can be used with atomic feature values (i.e. in a ‘flat’ structure) where the syncretism is the default to the ‘elsewhere’. In such cases, in the absence of other representational means, only one syncretic pattern can be described for any domain. Underspecification in such instances need not be equated with semantic naturalness, as the elsewhere is the residue of what is not specified in the morphology. Where a hierarchical feature structure is assumed, underspecification may represent the resort to an intermediate ‘natural class’, but as we saw in our discussion of Gaagudju in Chapter 4: §4.2.1, sometimes it is not possible to order feature values in such a way that we can always treat syncretism as a default to an intermediate semantic class.

Our case studies in §5.2–§5.4 concentrate on significant fragments of individual languages. In §5.2 we show that the Dhasaanac data in

Chapter 3: §3.7.2 can be analysed straightforwardly in the Network Morphology framework. In §5.3 we turn to the verbal system of Dalabon to show that syncretism which may originally have had a semantic or pragmatic motivation has been morphologized, and demonstrate that an adequate account requires a combination of referral and underspecification in order to see how that language's morphological systematicity relates to more general patterns. In §5.4 we look at the nominal system of Russian to show how the different degrees of morphological regularity can be represented in terms of the domains to which they apply, and how morphology, syntax and semantics interact in the animacy-based syncretisms. In each case, we have implemented our analysis of the languages in question, an important step, because it allows the reader to check that the analyses are valid, and so to have a firm basis for evaluating them.

5.1 Network Morphology and syncretism

We shall express our analyses within the Network Morphology framework. It is based on the following principles:

- (1) Morphology is lexeme-based.

This means that lexical items in a Network Morphology analysis are treated as generalizations over their inflectional paradigms.¹ Informally, the rules of morphology are therefore functions from lexical items into sets of inflected forms. A second important principle associated with Network Morphology is (2).

- (2) Morphology is inferential-realizational.

Network Morphology does not treat affixes (or other inflectional operations) as lexical entries. Instead these are represented as generalizations in an inheritance hierarchy. This is a natural way to treat inflection, as it is therefore straightforward to treat exceptionality and irregularity as associated with the lexeme in question, rather than having to encode in each affix the information about the most irregular item that it can combine with. We illustrate this further in §5.1.1.

¹ We are excluding from this definition idioms and clichés (such as 'throw down the gauntlet') which are listed lexically. However, even for items such as these it is still important to provide their inflectional variants (such as 'John threw down the gauntlet', and so on).

Another property of Network Morphology is that it uses default inheritance networks which can distribute information about morphology in different parts of the network.

- (3) Network Morphology uses default inheritance.

We illustrate how default inheritance works in §5.1.2. The main advantage of default inheritance theories is that one is able to specify the general inflectional properties of a language and at the same time include exceptions. It also allows one to see the domains over which a particular phenomenon applies. We shall illustrate this in our case study on Russian, where we consider the domains of syncretism in that language.

As with all formal frameworks we make a distinction between the fully specified morphosyntactic paradigm and the particular formal analysis which accounts for it. The fully specified morphosyntactic paradigm has the following characteristics:

- (4) Feature values are atomic.
- (5) Features are ordered relative to each other.²

When we talk of underspecification we mean the relationship of the formal analysis to the fully specified morphosyntactic paradigm.

- (6) The formal analysis may be underspecified relative to the fully specified paradigms.

Any underspecification of the formal analysis in (6) will have to obey the constraints of the ordering of features in (5). The ordering of the features reflects cross-linguistic tendencies, such as those illustrated in our analysis in Chapter 3: §3.8. Network Morphology uses underspecification and referrals and combines both as generalized referrals.

- (7) Generalized referral:
 - a. One feature specification may refer to another feature specification for its realization.

² There are at least two, logically independent, roles which might be attributed to feature ordering: (i) underspecification-based syncretism; (ii) the ordering of inflectional elements. The ordering of features, as we have seen in our discussion of uninflectedness and neutralization in Chapter 2 and the analysis in §3.8, is ideal for underspecification-based syncretism. We are not claiming that all rules of inflectional morphology can be accommodated with ordered features. Finkel, Shen, Stump and Thesayi (2002) develop KATR a multiset-based extension of DATR in order to treat instances of morphology for which feature ordering is not relevant.

- b. As with other realization rules, referrals may be underspecified.
- c. Extensions of the referring specification will be realized by extensions of the referred-to specification.

The consequence of (7a) is that there can be instances of directional syncretism. It follows from (7b) and (7c) that this directionality need not be limited to individual paradigm cells, but, because referrals are underspecified, may involve whole sub-paradigms.

While underspecification and generalized referrals, as we have formulated them, still involve reference to features, we allow for morphomic indexes to capture non-directional patternings which are independent of features. The relationship between these indexes and morphosyntax is still constrained by the ordering of features. In other words, indexed forms must still map to well-formed feature specifications.

- (8) Morphomic indexes must be associated with well-formed feature specifications.

We shall see how indexes are used in §5.2 in the case study for Dhaasanac.

An important part of the Network Morphology philosophy is that the analyses should be implementable and testable. For this purpose the lexical knowledge representation language DATR is used (Evans and Gazdar 1996). An attractive property of DATR is that it can be used to represent default inheritance networks, a key organizing principle of the Network Morphology framework.³

Default inheritance provides for elegant treatments of many problems within morphology. Corbett and Fraser (1993) and Fraser and Corbett (1995) showed how default inheritance could be used to analyse animacy, gender and declensional class assignment in Russian, and this work was applied to a substantial fragment of Russian nominal morphology. Brown (1998a) adopted this Network Morphology approach for analysing gender assignment in Polish, and Fraser and Corbett (1997) implemented a Network Morphology analysis of gender and morphological class assignment in Arapesh, a language of the Torricelli family, spoken on the north

³ DATR has been used for lexical knowledge representation for a variety of languages. The DATR-based work on German by Bleiching, Drexel and Gibbon (1996) and Cahill and Gazdar (1997, 1999) is of particular relevance for syncretism. Cahill and Gazdar (1997: 220–3) discuss rules of exponence and rules of referral in their work on adjectives, determiners and pronouns. This work is extended to nouns in Cahill and Gazdar (1999).

coast of Papua New Guinea. Evans, Brown and Corbett (2002) give a detailed account of the gender and morphological class assignment system of Bininj Gun-Wok (a non-Pama-Nyungan language of northern Australia, discussed above in Chapter 3: §3.5.1). This work included the mother-in-law, or avoidance register, and was applied to a sample of nouns in the language. The notion of layered defaults was important for that analysis, as certain nouns require access to what generally holds for their class and to what is the overall default for nouns as a whole. Other work has shown how the stress system of Russian can be analysed using default inheritance (Brown, Corbett, Fraser, Hippisley and Timberlake 1996), and Hippisley (1997, 2001) has applied the Network Morphology framework to word-formation.

5.1.1 *Inferential-realizational theories and morphology*

Network Morphology is an *inferential-realizational* theoretical framework. This term is due to Stump (2001: 1–30), who develops a typology of morphological theories which divides them according to two dimensions: *lexical* versus *inferential*, and *incremental* versus *realizational*. Lexical theories treat morphological formants, such as affixes, in the same way as stems and their associated meanings, by listing both in the lexicon. Inferential theories, on the other hand, treat the relation between lexical stems and their paradigm of word forms as a matter of inference, that is of rules or formulas. Incremental theories require of inflectional morphology that it be information-increasing so that words build their morphosyntactic specification as the result of acquiring affixes. Realizational theories, on the other hand, associate a set of morphosyntactic specifications with lexical stems and these specifications license morphological exponents, such as affixes.

Stump (2001: 17–27) gives a number of fundamental reasons for preferring inferential-realizational theories over other theories of inflectional morphology. What concerns us here are the specific properties of inferential-realizational theories which make them particularly well suited for analysing syncretism. Consider the Slovene noun paradigm in (9), reproduced from Chapter 4: §4.4.4, which has been discussed by Corbett and Fraser (1997) in relation to rules of referral.

The question here is how to guarantee that the plural stem *ljud-* occurs in the correct contexts in the dual, namely genitive and locative, and not in the other cases of that paradigm. This is a problem for an incremental framework, because a narrow feature specification for the stem is the most natural way of specifying the contexts in which they occur, but this would stop the addition of the affixes *-i* and *-eh*, because they would not

- (9) Paradigm of the Slovene noun *človek* 'person' (based on Priestly 1993: 401)

	singular	dual	plural
NOM	človek	človeka	ljudje
ACC	človeka	človeka	ljudi
GEN	človeka	ljudi	ljudi
LOC	človeku	ljudeh	ljudeh
DAT	človeku	človekoma	ljudem
INS	človekom	človekoma	ljudmi

contribute any new information and would therefore not be information-increasing. Lexical theories rely on the subcategorization properties of the affixes which they treat as having lexical entries. The combined suppletion and syncretism in (9) poses a number of problems for such theories. The stem *ljud-* requires a different set of plural affixes from those which would typically go with a regular noun of *človek*'s type. As we can see from the other cells in the dual paradigm, the stem *človek-* can be used in the dual. The question is what blocks the expected genitive plural/dual and locative plural/dual affixes for nouns of *človek*'s type, *-ov* and *-ih* respectively, from combining with the stem *človek-*? The answer for the plural paradigm is clear enough; there is a lexically specified suppletive plural stem, which is associated with a different set of endings. So the subcategorization information of affixes must be sensitive to the number information on the stem, even though most nouns do not have different singular and plural stems. At this point, a lexical framework already has to start listing within its regular affixes elements of potential exceptionality. For inferential-realizational theories this problem does not arise, because they do not treat affixal exponents as lexical entries that require exhaustive information on what they may combine with. In fact, inferential-realizational theories can treat exceptionality where it arises, with the stems of the lexical items in question, for which the appropriate affixes or other exponents can be inferred. When we continue to the more challenging question of the dual paradigm, the problems for lexical theories increase. If we assume that the stem *ljud-* is specified by some disjunction as also being the genitive dual and locative dual stem, then the affixes have to be sensitive to this information. If we do this, of course, then we have encoded in the affixes information about the most irregular items they can combine with. If our framework is a lexical-incremental one then things cannot be

worse, because the output information is no more informative than the subcategorization information. Yet, there is a simple answer to why we obtain the pattern in (9): for all nouns the genitive and locative dual and plural are always identical in form. Because they can make reference to sets of paradigm cells, inferential-realizational theories have no problem making simple generalizations like this.

While Network Morphology is an inferential-realizational framework, it also combines this property with the ability to incorporate varying degrees of regularity, by analysing morphology in terms of default inheritance, to which we now turn.

5.1.2 Default inheritance

Networks consist of *nodes* and connections between them. In default inheritance networks, information is inherited from higher nodes unless it is specifically overridden. In Figure 3 we present a simple default inheritance network. The diagram is based on a DATR example from Evans and Gazdar (1996: 176).⁴ In Figure 3, VERB, EN_VERB, Love, Do, Mow, Sew and Be are all nodes in the network. The nodes Love, Do, and EN_VERB inherit from VERB. The nodes Mow, Sew and Be inherit from EN_VERB, and therefore also from VERB. The nodes in Figure 3 are locations for information about the morphological paradigms of the items in question. As such, they generalize the information which classes of lexemes have in common. The relationships between nodes also make it possible to characterize the degree of exceptionality or lexical idiosyncrasy involved.

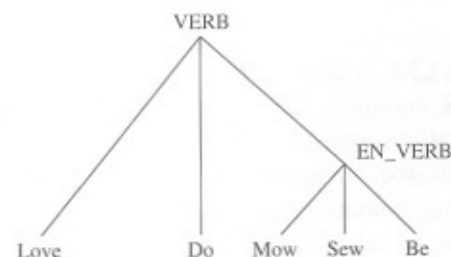


Figure 3. A default inheritance hierarchy for some English verbs (based on Evans and Gazdar 1996)

⁴ This section discusses default inheritance in DATR. For the purposes of illustration using English, the examples are based on the fragment published in Evans and Gazdar (1996), including their use of attributes. This fragment was written to demonstrate DATR, and not Network Morphology.

A property that all English verbs share is that the passive participle has the same form as the past participle. Examples (10) and (12) involve a past participle, while examples (11) and (13) involve passive participles.

- (10) I have *loved*.
 (11) I am *loved*.
 (12) I have *done* this.
 (13) This is often *done*.

We should be able to state this fact as a general property of verbs. The form of the past tense is generally *-ed*, but this can be overridden by particular items, such as *do*, whose past tense is *did*. Often, but not always, the past participle will have the same form as the past tense. There are also subregular classes, such as the one where the past participle is formed using *-en*. Default inheritance allows for a concise treatment of these facts. Evans and Gazdar (1996: 176) state the following at the node VERB in (14), where we have omitted some information, as indicated by the ellipses. What is given in (14) is a representation of the information associated with the top node in Figure 3.

- (14) VERB:
 <syn cat> == verb
 <syn type> == main
 <mor past> == "<mor root>" ed
 <mor passive> == "<mor past>"
 <mor present> == "<mor root>"
 <mor present participle> == "<mor root>" ing
 <mor present tense sing three> == "<mor root>" s
 ...

The node name VERB is placed before the colon. Each line containing '=' is a DATR equation. Each left-hand side of a DATR equation contains *paths*. Paths contain a combination of ordered *attributes*. The right-hand side of the equation may contain *values*, such as 'verb'. Alternatively it may contain paths, or node names, or it may contain a combination of paths, values and node names.

The first equation at VERB states that the syntactic category of items belonging to this class is 'verb'. The equation after this states that the syntactic type of verb is 'main' (i.e. a typical verb is a main verb rather than an auxiliary). The next equation says that the past is a concatenation of *-ed* onto what Evans and Gazdar call the morphological root. The equation after that says that the passive has the same form as the past. This is the way referrals are represented in DATR. The equation after the

statement about the passive says that the present uses the morphological root. This brings us to another important property of DATR: in the absence of any information to the contrary, we can infer that the value of a path will be the same as the value for the most specific path of which that path is an extension. In (15) we give examples of path extension.

- (15) <mor present>
 <mor present participle>
 <mor present tense sing one>
 <mor present tense sing two>
 <mor present tense sing three>
 <mor present tense plur>

All of the paths listed after <mor present> in (15) are extensions of it. At the node VERB in (14) only two of these extensions are found on the left-hand side of equations, in addition to the path <mor present> itself. These are:

- (16) <mor present participle>
 <mor present tense sing three>

This means that we can infer that the values for the first person singular present tense, the second person singular present tense and the plural present tense are the same as for the present tense as a whole, namely the morphological root. This is an example of *default inference*. The values for the present participle and third person present tense are not inferred in this way, as they are already specified at VERB.

From the equation at VERB with left-hand side <mor past> it can be inferred that the past tense and past participle (which are extensions of <mor past>) are both formed by suffixing *-ed* to the morphological root. We have already seen a referral-based way of stating the identity of the past participle and passive participle. Here the default syncretism of the past tense and past participle is the result of underspecification.

We also need to override the default syncretism of the past tense and past participle. Verbs which have a past participle in *-en* must inherit the suffixation of *-ed* for the past tense but override it for the past participle. This can be stated at the node EN_VERB which, as we can see from Figure 3, inherits from the node VERB. Evans and Gazdar (1996: 176) represent this in DATR as follows:

- (17) EN_VERB:
 <> == VERB
 <mor past participle> == "<mor root>" en.

In (17) the path <> indicates that EN_VERB will inherit all information from VERB unless otherwise specified. EN_VERB therefore inherits the specification that <mor past> suffixes *-ed* onto the morphological root. This means that the past tense forms for EN_VERB will still be formed using *-ed*. As the extension <mor past participle> has a specific value at EN_VERB, this means that the past participles of verbs of this type will suffix *-en*. This is one way in which the default syncretism of the participle and the past tense forms at VERB is overridden by a particular class of verbs, and it is represented in terms of path extension, which can be interpreted as being a strict form of underspecification. It is strict, as the order of attributes in the path is important for the definition of extension. (17) also illustrates another point about theories which use default inheritance, namely that they can involve layered or cascaded defaults. The overall default for verbs is to have the past participle and past tense form the same, but another layer of verbs maintains the default *-ed* for the past tense, while introducing its own default for the past participle.

Syncretism can be represented in terms of underspecification and referral in Network Morphology. An important advantage of this is that it is possible to combine referrals and underspecification to pick out sets of cells which are syncretic. In the following sections we discuss how underspecification relates to semantic naturalness and how to treat systematic syncretisms. The beauty of an inheritance-based approach, such as Network Morphology, is that it can represent what occurs over a large domain of morphology, representing degrees of regularity, from the domain of a particular word class, or higher, right down to individual lexical items.

5.1.3 *Underspecification and semantic naturalness*

Underspecification is often associated with a 'semantic' approach to syncretism. However, the feature values involved in syncretism may or may not be semantically assigned. This is true of gender, for example, in systems which combine formal and semantic assignment. Under such systems the relationship between genders *x* and *y* might be formal (in that they share some morphology), and between genders *y* and *z* semantic (in that they are related by meaning). Or genders *x*, *y* and *z* may have a semantic core complemented by formal assignment for each. Given the cross-linguistic propensity of genders to collapse (Chapter 3: §3.8), particularly when number is involved (Chapter 3: §3.4.2), we might also wish to treat gender

syncretism as underspecification for this reason. This means that there are therefore two different interpretations of what underlies underspecification:

- (18) Syncretisms based on underspecification reflect semantic naturalness.
- (19) Syncretisms based on underspecification reflect the cross-linguistic tendency of a feature to syncretize (i.e. for different values of that feature to share identical morphology).

The more appropriate interpretation of underspecification is that in (19). The reason for this is that underspecification is the natural way of treating both uninflectedness and neutralization, as we saw in Chapter 2: §2.4. If we consider the examples of neutralization and uninflectedness given in Chapter 2: §2.4, none of them can be explained entirely in terms of the structure inherent to the feature involved. As we have seen in Chapter 3: §3.4.2, there appears to be some involvement of number – typically plural – and the restructuring of gender, which may sometimes have a semantic basis. Uninflectedness is a morphological property which is sensitive to morphosyntactic structure. Under such circumstances underspecification should be put in the context of other morphosyntactic features and is itself not directly interpretable as reflecting a semantic motivation for syncretism.

It is logical to ask whether syncretism can be determined solely by semantics. Animacy in Slavonic appeared to be a possible instance. But as we saw in Chapter 3: §3.1, it is difficult to isolate an on-going semantic pattern underlying the instances of type 2 syncretism (core case with non-core case), to which this Slavonic example belongs. Furthermore, the animacy rule is itself subject to being overridden when items which should be subject to it belong to a particular inflectional class, showing that morphology plays a role.⁵ As Claire Bower points out (personal communication) there are instances in various other languages of place names having their nominative form identical to the locative. However, these are typically instances of lexically determined syncretism: individual items have the syncretism, but not all the items in a semantic class. In Tsez, proper nouns denoting places, provided they are native words, will have the absolutive identical to one of six local cases (Bernard Comrie, personal communication). For example, the name of the village *Asakh*, *asaq*, includes the marker *-q* 'on (a vertical surface)' and may function both as

⁵ The noun *djadja* 'uncle' in Russian, for example, is a masculine animate noun but is not subject to the animacy rule in the singular, because of its inflectional class membership.

an absolutive and as a local case form. Hence the semantic type allows the prediction that there will be a syncretism but does not determine which case will be involved, since this varies from place name to place name. Moreover, though the form may be used as an absolutive, speakers show some reluctance here, and prefer to use the name in apposition to a common noun, which has a clear absolutive. Here again, then, semantics does not uniquely determine a domain of syncretism. To date we have not found a language in which a noun denoting a place will necessarily have a specific syncretism. And in general, we have not found instances of semantics uniquely providing a domain for syncretism.

5.1.4 *Systematicity in Network Morphology*

Below we offer a model in which systematic syncretism is represented by the devices of underspecification, indexing and referral, supplemented by the DATR notion of attribute ordering.

In the analysis which we are proposing underspecification is related to the ordering of attributes in a path, which in turn is connected with the cross-linguistic patterns of ordering of features which were highlighted in Chapter 3: §3.8. Hierarchies which rank inflectional features have played an important role in other accounts of particular areas of the syncretism spectrum: for example, Carstairs (1987) treatment of 'take-overs' and the Systematic Homonymy Claim as discussed in Chapter 4: §4.4.2, or Stump's (2001: 239) Feature Ranking Principle, as discussed in Chapter 4: §4.4.4. Our formal analysis is based on the following interpretations:

(20) *Network Morphology analysis*

- a. The order of attributes in a path is associated with cross-linguistic tendencies of particular features to syncretize. The further to the right an attribute occurs, the more likely it is to syncretize.
- b. The order of attributes in a path will line up with neutralization and uninflectedness.
- c. Underspecification involves resort to an elsewhere form only.
- d. Referrals represent directional systematic syncretism within the language.
- e. Indexes represent non-directional syncretism within the language.
- f. Syncretisms which violate the order of attributes as required by (20a) are language-specific referrals (but still systematic).
- g. For underspecification-based syncretism there is generally one form which has more than one function as a default property.

- h. For referral-based syncretism there is a default form associated with a default primary function. The other function of that form will typically be manifested in a subset of lexical items or morphosyntactic contexts.
- i. The adequate treatment of some syncretic patterns requires simultaneous use of underspecification and referrals, thereby justifying the theoretical requirement for both. The Slovene nominal paradigm in (9) and the Dalabon verbal system (§5.3) involve examples of this type.

The order of attributes in (20a) is established according to cross-linguistic tendencies, such as those observed in Chapter 3: §3.8. Given atomic values, we can easily find examples of less straightforward syncretism which line up with uninflectedness and neutralization, as expected by (20b). Breaking features up further may obscure this tendency, as we will allow for explanations based more on uninflectedness and neutralization. Because of the constraint in (20a) and the requirement for atomic values in (5), the number of possible elsewhere forms is constrained by the number of features realized. For example, if we considered a language which has two number values and three or more case values, then it potentially has two elsewhere forms, a singular one and a plural one. This arises from ordering number before case (as suggested by the data in Chapter 3: §3.8). In contrast with this, the potential number of elsewhere forms for Russian adjective morphology is twelve (a grid of two number values and six case values), if number is ordered before case, and case is ordered before gender. If we ignore animacy, then Russian comes close to making use of the full range of elsewhere forms in the adjectival paradigm: syncretism of masculine and neuter in the four oblique cases in the singular, and neutralization of all gender distinctions in the plural, leaving just six plural cells. But this is not the end of syncretism in Russian, because there are additional patterns related to animacy.

Referral-based syncretism is not constrained by an underlying paradigmatic design which leads to a limit on the number of elsewhere forms. However, we would also expect to find some indication of directional effects.

5.2 *Case study 1: Dhaasanac*

As we saw in Chapter 3: §3.7.2 the Dhaasanac verbal paradigm is an example of morphological regularity where the syncretized values are

heterogeneous. This is a very good example of 'morphology by itself' (Aronoff 1994), as the use of indexes below shows. In the case study we show how we can capture this autonomous morphological regularity and still associate it with the relevant morphosyntax. The DATR representation associated with this analysis is given in Appendix 4. We repeat the scheme of syncretism in Dhaasanac from Chapter 3: §3.7.2 in (21).

(21) Dhaasanac syncretisms

	singular	plural
1 INCL	—	A
1	A	B
2	B	B
3 F	B	A
3 M	A	A

This pattern generalizes across the positive imperfect, positive perfect, positive dependent and positive simple past paradigms in Dhaasanac. This is strong evidence for morphological systematicity, which is further reinforced by the fact that different verbs have different A and B variants but still obey the pattern in (21).

The other important aspect of the Dhaasanac data is that they are an instance of mediated polarity, a term which we introduced in Chapter 3: §3.7.1. Importantly, the B form is always used for the second person, both singular and plural. This is therefore an example of underspecification of number in the presence of second person, but which still needs to be combined with indexing of the forms. The fully specified paradigms which we assume for Dhaasanac will be as in (22). Here 'pos' stands for positive, 'prf' for perfect, and '1st sg' for first person singular, and so on.

(22) <mor pos prf 1st sg>
 <mor pos prf 2nd sg>
 <mor pos prf 3rd sg f>
 <mor pos prf 3rd sg m>
 <mor pos prf 1st_excl pl>
 <mor pos prf 1st_incl pl>
 <mor pos prf 2nd pl>
 <mor pos prf 3rd pl>

The ordering of attributes in (22) reflects the fact that it is negation (or its absence) and TAM which determine whether there are A and B forms,

because the negated paradigms and the imperative do not distinguish A and B forms. It is the role of our formal analysis to infer the correct forms for the positive paradigms.

In (23) we give the lexical entry for the verb *fúr* 'to open', which belongs to the class of verbs whose stem ends in a coronal consonant.

(23) Fúr:
 <> == VERB
 <form> == CORONALS
 <gloss> == 'open'
 <stem type> == one_vowel
 <stem cons_1> == f
 <stem vowel_1> == _u
 <stem final> == _r.

The node Fúr inherits from the node VERB, which includes the information in (24). (We have omitted some of the information given at VERB.)

(24) VERB:
 <> ==
 <syn> == verb
 <index> == _A
 <index 2nd> == _B
 <index 3rd sg f> == _B
 <index 1st_excl pl> == _B
 <mor pos imprf> == "<form imprf <index> >"
 <mor pos prf> == "<form prf <index> >"
 ...

Recall from (22) that extensions of the path <mor pos prf>, for example, will involve attributes for person and number (in that order), and for the third person singular, also gender. Hence, in the absence of any more specific information, the form of the perfect paradigm will be determined by looking at the equation associated at VERB with <mor pos prf>, namely <mor pos prf> == "<form prf <index> >". The right-hand side of the equation "<form prf <index> >" involves an evaluable path (the <index> part). This equation basically means that the morphology of the positive perfect is determined by looking for the form of the perfect and inserting the appropriate index for that form. Let us consider the morphology for the path <mor pos prf 2nd sg>. The rules of inference mean that, among other things, we can infer (25) from the equation <mor pos prf> == "<form prf <index> >".

(25) We can infer that:
 <mor pos prf 2nd sg> == "<form prf <index 2nd sg> >"

From the information in (24) we can infer (25), and because of the equation $\langle \text{index } 2\text{nd} \rangle == _B$ in (24), and the inference made in (25), we can infer (26).

(26) $\langle \text{mor pos prf } 2\text{nd sg} \rangle == \langle \text{form prf } _B \rangle$

The other TAM, person and number combinations work in the same way. This is a realizational approach to morphology because it separates out the morphosyntactic specification from the actual form with which it may be associated. We should note from (23) that the lexical item *Fúr* inherits information about its associated forms from the node CORONALS. Even though *Fúr*'s A and B forms may differ from those of other verbs, the equations at VERB can generalize across all verbs, irrespective of the specific realization of the A and B forms. In the case of *Fúr* we can obtain the forms in (27) for the perfect paradigm.

(27) *Fúr*: $\langle \text{mor pos prf } 1\text{st sg} \rangle = f_u_r-i$.
Fúr: $\langle \text{mor pos prf } 2\text{nd sg} \rangle = f_u_d-d-i$.
Fúr: $\langle \text{mor pos prf } 3\text{rd sg f} \rangle = f_u_d-d-i$.
Fúr: $\langle \text{mor pos prf } 3\text{rd sg m} \rangle = f_u_r-i$.
Fúr: $\langle \text{mor pos prf } 1\text{st_excl pl} \rangle = f_u_d-d-i$.
Fúr: $\langle \text{mor pos prf } 1\text{st_incl pl} \rangle = f_u_r-i$.
Fúr: $\langle \text{mor pos prf } 2\text{nd pl} \rangle = f_u_d-d-i$.
Fúr: $\langle \text{mor pos prf } 3\text{rd pl} \rangle = f_u_r-i$.

Thus, by treating the inventory of forms available as separate from the associated morphosyntax, it is possible to capture the morphological systematicity of the Dhasaanac verb. At the same time the morphology is still constrained, because these indexed items have to match up with well-formed feature specifications.

5.3 Case study 2: The Dalabon verbal system

This case study sets out the core details of syncretism within Dalabon. It is based on Evans, Brown and Corbett (2001) and uses the same orthography. The DATR representation associated with this Network Morphology analysis is given in Appendix 5.

Dalabon is an Australian language of the Gunwinyguan family, spoken in central Arnhem Land by a declining population which is now reduced to about twenty fluent speakers. Dalabon's verbal system marks intransitive subjects using bound prefixal morphology, and transitive subjects and objects using a combination of bound prefixes and clitics. As we saw in

Chapter 3: §3.3 there is strong cross-linguistic evidence to show that transitive verbs pattern very differently from intransitives in terms of the person syncretism. This would suggest that transitivity is an important factor in syncretism on verbs.

Dalabon's verbal system involves a morphologization of what was originally a semantically or pragmatically motivated avoidance strategy, a phenomenon investigated by Heath for Australian languages (Heath 1991) and languages of the Americas (Heath 1998). The avoidance strategy in Dalabon has produced a situation in which certain combinations of first person subject on second person object or second person subject on first person object are avoided by the use of person syncretism of the subject, whereby the appropriate form of the third person subject is used. There is a directional effect, because the forms used in the syncretism have third person subject as their primary function, as can be demonstrated by looking at the intransitive paradigm. Furthermore, the referrals are simultaneously combined with underspecification and predict whole sets of cells in the extended Dalabon paradigm. This demonstrates that the best treatment of some syncretisms requires a combination of referral and underspecification, and so an adequate framework requires both.

5.3.1 *The structure of the Dalabon intransitive paradigm*

Dalabon pronominal prefixes mark tense, aspect, mood and clause status. There are six relevant distinctions: REALIS, SUBORD1, SUBORD2, IRREALIS, APPREHENSIVE and HORTATIVE. With the exception of certain forms in the APPREHENSIVE, the six series for tense, aspect, mood and clause status can be derived by regular rules. The other series can be derived from the SUBORD1 forms. The prefix forms shown in this case study are the REALIS prefixes, which differ from the SUBORD1 forms in that they terminate with a marker *-h*.⁶

The forms linked to certain cells within the transitive paradigm are based on the corresponding intransitive ones, and so we look at the features of the intransitive paradigm first. We represent the features involved as a 4×4 matrix, with four person values (FIRST EXCLUSIVE, FIRST INCLUSIVE, SECOND and THIRD), combined with three values for the number feature: SINGULAR,

⁶ The values used here are based on those in Evans, Brown and Corbett (2001). Evans has recently revised the inventory of values, merging SUBORD1 and IRREALIS, and adding a further value, PURPOSIVE (Nicholas Evans, personal communication). As the relevant issue for our case study is person syncretism which generalizes across the tense, aspect, mood and clause status series, this change is not significant for the formal analysis presented here.

DUAL, PLURAL, with an additional feature of harmonicity which has two values, HARMONIC and DISHARMONIC. This feature marks kinship relations and is normally restricted in its interpretation to the subjects of intransitive and transitive verbs, but may occasionally be interpreted in relation to the object (see Evans, Brown and Corbett 2001: 197–8). Typically, harmonicity can be seen as a distinction made in the presence of dual number.

The Dalabon intransitive paradigm (REALIS) is given in (28).

(28) The Dalabon intransitive paradigm (REALIS series)

	I EXCL	I INCL	2	3
SG	ngah-	—	djah-	kah-
DU DIS	ngeh-	djeh-	deh-	keh-
DU HARM	yarrah-	yah-	narrah-	barrah-
PL	yalah-	ngarrah-	nalah-	balah-

Logically there cannot be a first person inclusive singular, hence the gap in the paradigm in (28). From this it follows that person in Dalabon determines number marking to an extent. We can therefore represent the relationship between person, number and harmonicity in Dalabon as a hierarchy in Figure 4. The ordering of attributes in the paths directly reflects the hierarchy in Figure 4. The hierarchy in Figure 4 is therefore a representation of the structure of the verbal paradigm for which our formal analysis should predict the forms. The FIRST INCLUSIVE cannot have singular forms, and so the ordering allows us to state that there are no SINGULAR extensions of FIRST INCLUSIVE. It is important to make a distinction here between the fully specified paradigm in Figure 4, on the one hand, and the statements required in the formal analysis to infer the correct forms in that fully specified paradigm. When we turn to the structure of the transitive paradigm in the next section, we will see that we need to refer one person value to another, together with its possible extensions for number. This illustrates the point that referrals and underspecification are required simultaneously.

5.3.2 *The Dalabon transitive paradigm and the inadequacy of underspecification*

There are 102 distinct subject/object combinations, which in turn generate $102 \times 6 = 612$ forms when all TAM combinations are included. This

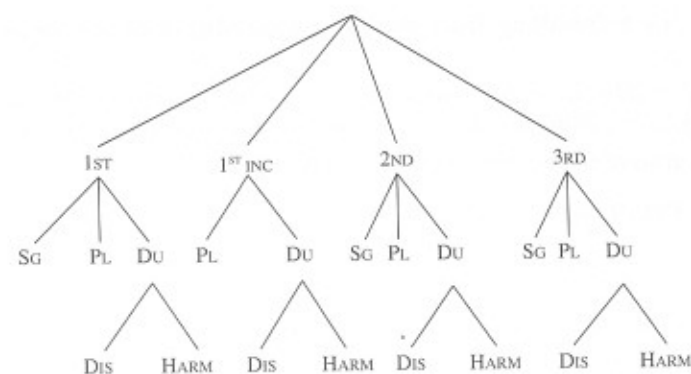


Figure 4. Paths for person, number and harmonicity in Dalabon

number takes account of the fact that coreferential combinations are ruled out, as these are encoded by the use of a reflexive/reciprocal suffix with the intransitive paradigm. There is also a contrast between 'higher' and 'lower' third person objects when both subject and object are third person singular. We shall analyse the paradigm section by section.

In (29) we give the forms of the paradigm for the third person transitive subject, when the object is singular. For comparison, the intransitive subject forms are given in square brackets above the columns with the transitive forms. The forms for the second person singular object are

(29) The third person REALIS subject paradigm in Dalabon

		subject			
		3SG	3 DIS	3DU	3PL
object	[none	kah-	keh-	barrah-	balah-]
	1SG	kah-	keh-	burrah-	bulah-
	2SG	djah-	djirrah-	djirrah-	djilah-
	3SG	kah-	keh-	burrah-	bulah-
		bvkah ⁷ -			

special portmanteau forms which combine information about the subject and the object. In the case of the paradigm for the second person singular object it is possible to interpret the syncretism of 3 DIS > 2SG with

⁷ Used for a higher animacy object.

3DU > 2SG as resulting from the underspecification of the values for harmonicity.

The disyllabic dual and plural transitive subject forms which occur in (29) with the first person objects and third person objects are derived from the intransitive subject forms by the following rule:

- (30) CaLah- → CiLah- if C = y
 →CuLah- elsewhere
 (where L is a liquid, i.e. l or rr)

The forms derived by rule (30) could be interpreted as marking transitivity, subject person and number, while being underspecified for the actual person and number of the object. This is because they can be used with a third or first person singular object or with non-singular objects of any person. In the latter case a variety of proclitics are used to mark the object.

The analysis of the forms *burrah-* and *bulah-* has consequences for our analysis of the forms involving the second person singular object. If *djirrah-* and *djilah-* are treated as realizing second singular object combined with the number of the subject, then the competition between the forms *burrah-* and *djirrah-*, and between the forms *bulah-* and *djilah-*, cannot be resolved using Panini's principle, as explained in Chapter 4: §4.3, because the feature specifications would overlap but neither would be more specific than the other (31).

- (31) *burrah-* (A 3DU), *bulah-* (A 3PL)
djirrah- (OBJ 2SG, A DU), *djilah-* (OBJ 2SG, A PL)

The reason that the competition cannot be resolved using Panini's principle is that the forms *burrah-* and *bulah-* in (31) contain a subject feature value for third person.

- (32) CONFLICT A: which forms should be used for 3DU > 2SG and 3PL > 2SG?

However, we can avoid CONFLICT A, if the value for the subject person in *burrah-* and *bulah-* is deleted, treating these prefixes as marking only subject number. We would then obtain (33).

- (33) *burrah-* (A DU), *bulah-* (A PL)
djirrah- (OBJ 2SG, A DU), *djilah-* (OBJ 2SG, A PL)

Now the specification for *burrah-* and *bulah-* is contained in that for *djirrah-* and *djilah-* and so the competition between the two can be resolved according to Panini's principle. So far it looks as though the specifications in (33) will obtain the right result. This appears to be borne out when we

look at other parts of the transitive paradigm, such as that for the first exclusive subject in (34).

- (34) The first person exclusive REALIS subject paradigm in Dalabon

		subject			
		1SG	1DIS	1DU	1PL
object	[none	<i>ngah-</i>	<i>ngeh-</i>	<i>yarrah-</i>	<i>yalah-</i>
	2SG	<i>djah-</i>	<i>djirrah-</i>	<i>djirrah-</i>	<i>djilah-</i>
	3SG	<i>ngah-</i>	<i>ngeh-</i>	<i>yirrah-</i>	<i>yilah-</i>

The forms for 1DU > 3SG and 1PL > 3SG are obtained from the intransitive subject forms by the rule in (30). This means that we can assume that they have the specifications in (35).

- (35) *yirrah-* (A 1DU), *yilah-* (A 1PL)

Recall from (33) the specifications for *djirrah-* (OBJ 2SG, A DU) and *djilah-* (OBJ 2SG, A PL). As things stand, the specifications for *djirrah-* and *djilah-* overlap with those for *yirrah-* and *yilah-* in (35), but neither is more specific than the other.

- (36) CONFLICT B: which forms should be used for 1DU > 2SG and 1PL > 2SG?

The natural resolution of CONFLICT A by underspecification in (32) was to treat the forms *burrah-* (A DU) and *bulah-* (A PL), which are used for 3DU > 2SG and 3PL > 2SG, as not marking subject person. This means that the option of underspecifying the subject person of *yirrah-* (A 1DU) and *yilah-* (A 1PL) is not open to us in order to resolve CONFLICT B in (36), as this would then be the same specification as for *burrah-* (A DU) and *bulah-* (A PL). Hence the resolution of CONFLICT A is itself in conflict with the resolution of CONFLICT B. We could claim that this problem can be avoided by resort to a principle which prioritizes the realization of the 2SG object. While the second person singular object is clearly being treated as special, the problem is that first person objects are not realized by bound prefixal morphology, and so the putative principle cannot be associated with a position to the left on a person hierarchy, and would therefore involve ad hoc stipulation.

There is a further problem for an underspecification analysis when we turn to the second person subject paradigm in (37).

(37) The second person REALIS subject paradigm in Dalabon

		subject			
		2SG	2DIS	2DU	2PL
object	[none	<i>djah-</i>	<i>deh-</i>	<i>nurrah-</i>	<i>nalah-</i>
	1SG	<i>kah-</i>	<i>keh-</i>	<i>burrah-</i>	<i>bulah-</i>
	3SG	<i>dah-</i>	<i>deh-</i>	<i>nurrah-</i>	<i>nulah-</i>

Like *yirrah-* and *yilah-* the forms *nurrah-* and *nulah-* are derived by rule (30) from the intransitive forms and therefore have the feature specifications in (38).

(38) *nurrah-* (A 2DU), *nulah-* (A 2PL)

However, the question arises why the forms *burrah-* and *bulah-* are used for the 2DU > 1SG and 2PL > 1SG respectively. To see this, let us compare the feature specifications with those of *burrah-* and *bulah-* which were given in (33).

(39) *nurrah-* and *nulah-* should fill the 2DU > 1SG and 2PL > 1SG cells according to Panini's principle

burrah- (A DU), *bulah-* (A PL)
nurrah- (2A DU) *nulah-* (A 2PL)

As the feature specifications in (38) for *nurrah-* (A 2DU) and *nulah-* (A 2PL) are more specific than the feature specifications we assumed for *burrah-* (A DU) and *bulah-* (A PL) in (33), we would expect *nurrah-* and *nulah-* to fill the 2DU > 1SG and 2PL > 1SG cells. However, even if we eliminated the person information of *nurrah-* and *nulah-*, this would not make them less specific than *burrah-* and *bulah-*. Hence, underspecification cannot avoid the incorrect prediction that *nurrah-* and *nulah-* will fill the 2DU > 1SG and 2PL > 1SG cells.

(40) CONFLICT C: which forms should be used for 2DU > 1SG and 2PL > 1SG?

One way to resolve CONFLICT C in (40) might be to alter the information content in (38) so that *nurrah-* and *nulah-* specify the person of the object, namely third person. However, a further problem, which is perhaps not insurmountable, is that the rule in (30) would then be associated with different morphosyntactic content, depending on the prefix. For

nurrah- and *nulah-* rule (30) would not only involve the addition of information that the prefix is transitive, but also that the person of the object is third person. On the other hand, for *yirrah-* (A 1DU) and *yilah-* (A 1PL) rule (30) would not involve the addition of information about the person of the object.

So far we have concentrated on the non-singular subject forms which are derived by rule (30) from their intransitive counterparts. But it turns out that the problems we have highlighted for underspecification and the non-singular forms generalize to all forms, that is to singular and disharmonic as well. In (30) we present the paradigm of singular subjects and singular objects.

(41) The paradigm of singular subjects and singular objects (REALIS)

		subject		
		1SG	2SG	3SG
object	[none	<i>ngah-</i>	<i>djah-</i>	<i>kah-</i>
	1SG	—	<i>kah-</i>	<i>kah-</i>
	2SG	<i>djah-</i>	—	<i>djah-</i>
	3SG	<i>ngah-</i>	<i>dah-</i>	<i>kah-</i>
		bvkah- (higher animacy object)		

For the singular subjects similar problems arise for underspecification. If *djah-* may mark the second singular intransitive subject and the object in the 1SG > 2SG cell (an ergative ~ absolutive pattern), then why is *kah-* used in the 2SG > 1SG cell, rather than *djah-*? We can, of course, specify *kah-* as marking a singular subject (with person underspecified), but even then *djah-*, which marks second person singular (subject or object), would be expected to fill the 2SG > 1SG cell, being more specific. This problem is in some senses worse than CONFLICT A, as in this case it is not that the feature specifications overlap, but that one is more specific than the other, and the wrong one (the less specific *kah-*) wins. The conflict between *djah-* (2SG) and *ngah-* (1SG) to realize the 1SG > 2SG cell can be seen as an extension of CONFLICT B. Again, we cannot underspecify *ngah-* for subject person, as this would then mean it had exactly the same specification as *kah-* (which is the same problem we encountered for *yirrah-* (A 1DU) and *yilah-* (A 1PL)), and here it is not clear that our putative principle, which

treats the marking of second person objects as special, can work, as *djah-*, because of its alignment, is not just an object marker. So for a variety of reasons underspecification fails when used on its own.⁸

5.3.3 The Dalabon transitive paradigm: a generalized referral analysis
As we showed in Chapter 3: §3.3, the person syncretisms which arise in transitive verbs are a case apart from intransitives. This suggests that, where we seek explanations for syncretisms in transitive verbs, these should be justified on the basis of combinations which are brought about by transitivity. At the same time we should not dispense with the insights of underspecification, if it can be applied naturally.

So far we have looked only at those parts of the Dalabon verbal paradigm involving singular objects. If we lay out the whole paradigm (Table 2), then a clear pattern emerges.

- (42) GENERALIZATION 1: The 1SG > 3SG, 3SG > 3SG, 1 DIS > 3SG and 3 DIS > 3SG forms are the same as the forms for the corresponding intransitive subject (with the exception of *bvka-* which marks 3SG > 3SG with higher animacy objects).
- (43) GENERALIZATION 2: The 1DU > 3SG, 3DU > 3SG, 1PL > 3SG, 3PL > 3SG forms are transparently derived from the intransitive subject forms by rule (30).

Generalizations 1 and 2 both have a phonological basis since the forms referred to in (42) are all monosyllabic, and those referred to in (43) are polysyllabic. It should also be noted that third person singular object marking cannot be treated as neutralization, as we have defined it in Chapter 2: §2.4.1, for two reasons. First, third person agreement occurs with a noun phrase, as in (44) where the comitative verb 'appear with' agrees with the noun *Naworneng*.

- (44) "djarra Naworneng nga-h-ye-burlhmung" kah-yini-nj
here Naworneng 1.SG>3.SG-R-COM-appear.PST.PFV 3.SG-say-PST.PFV
"Here I've brought Naworneng" he said.' (Evans, Brown and Corbett 2001: 200)

⁸ This conclusion finds independent justification in that Wunderlich (2001a) introduces 'taboos' as a constraint in his correspondence-theoretic account of Dalabon. The mechanisms involved in this analysis basically replace certain feature values with others by deleting particular values in the input. This is actually very similar to referrals, thereby demonstrating that underspecification alone cannot account for the Dalabon data.

Table 2: Paradigm of Dalabon subject + object combinations, including object clitics; realis TAM series (based on Evans, Brown and Corbett, 2001: 199).

subject [INTR /TR] form]	1 [ngah-]	2 [djab-] [A dab-]	3 [kah-]	1 DIS [ngeh-]	2 DIS [deh-]	3 DIS [keh-]	1DU [yirrah-] [A yirrah-]	2DU [nurrab-] [A nurrab-]	3DU [burrab-] [A burrah-]	1PL [yilab-] [A yilab-]	2PL [nulah-] [mulah-]	3PL [bulah-] [A bulah-]
1 [ngey]		kah-	kah-		keh-	keh-		burrab-	burrab-		bulah-	bulah-
2 [njing]	djab-		djab-	djirrah-	deh-	djirrah-	djirrah-	nurrab-	djirrah-	djilab-	nulah-	djilab-
3 [-]	ngah-	dah-	kah- BVKAH-	nggeh-	deh-	keh-	yirrah-	nurrab-	burrab-	yilab-	nulah-	bulah-
1DU [njerr]		njerr kah-	njerr kah-		njerr keh-	njerr keh-		njerr burrab-	njerr burrab-		njerr bulah-	njerr bulah-
2DU [norr]	norr ngah-		norr kah-	norr ngeh-		norr keh-	norr yirrah-		norr burrab-	norr yilab-		norr bulah-
3DU [bu(l)nu]	bunu ngah-	bunu dah-	bunu kah-	bunu ngeh-	bunu deh-	bunu keh-	bunu yirrah-	bunu nurrab-	bunu burrab-	bunu yilab-	bunu nulah-	bunu bulah-
1PL [njel]		njel kah-	njel kah-		njel keh-	njel keh-		njel burrab-	njel burrab-		njel bulah-	njel bulah-
2PL [nol]	nol ngah-		nol kah-	nol ngeh-		nol keh-	nol yirrah-		nol burrab-	nol yilab-		nol bulah-
3PL [bulu]	bulu ngah-	bulu dah-	bulu kah-	bulu ngeh-	bulu deh-	bulu keh-	bulu yirrah-	bulu nurrab-	bulu burrab-	bulu yilab-	bulu nulah-	bulu bulah-

The combinations with first inclusive have been omitted from this table (for which see Evans, Brown and Corbett 2001: 199). The prefixal marking in bold is determined by the referrals (56) and (57).

A second reason is the form *bykah-* which marks a higher animacy third person object. As this form is used uniquely for third person objects (in the presence of a third person subject), it suggests that third person is a valid syntactic distinction for objects in Dalabon. Both of these facts indicate that object person marking in the singular is not irrelevant for syntax, which therefore means that this cannot be seen as neutralization.

The next set of generalizations concerns the first person singular object.

- (45) GENERALIZATION 3: The 3SG > 1SG and 3DIS > 1SG forms are the same as the forms for the corresponding intransitive subject.
- (46) GENERALIZATION 4: The 3DU > 1SG and 3PL > 1SG forms are transparently derived from the intransitive subject by rule (30).
- (47) GENERALIZATION 5: Generalizations 3 and 4 mean that the forms of 3 > 1SG are the same as the forms of the 3 > 3SG paradigm (with the exception of *bykah-*).
- (48) GENERALIZATION 6: Generalizations 3 and 4 mean that there is no marking of first person singular objects in the bound pronominal morphology.

Generalizations 3–6 might lead one to assume that first person singular object is not a morphosyntactically relevant entity. However, if it played no role, or indeed were non-existent, then CONFLICT C in (40) could not arise. This point itself is independent of how we actually choose to treat CONFLICT C, by underspecification or otherwise. If, for example, first person object were a syntactically non-existent entity then we would have no need to refer to it in the paradigms and also in the formal analysis associated with them. But this would mean that the contrast between the forms *kah-* (2SG > 1SG) and *dah-* (2SG > 3SG) in (41) could not exist. Of course, the prefixes *kah-* and *dah-* are used for other purposes and would not be treated in most theories as fully specified for the feature values 2SG > 1SG and 2SG > 3SG respectively. Examination of Table 2 shows that there is always a contrast between 2 > 1 and the 2 > 3 paradigm, and this means that reference must be made by morphology to first person objects. As for the second person, we know that objecthood is important there, because of the special port-manteau forms.

Having demonstrated the inadequacy of underspecification alone, we now turn to the analysis in Evans, Brown and Corbett (2001) and our representation of it. That analysis resolves CONFLICT A by specifying the feature values for *burrah-*, *bulah-*, *djirrah-* and *djilah-* as in (49).

- (49) *burrah-* (A 3DU), *bulah-* (A 3PL)
djirrah- (OBJ 2SG, A 3DU), *djilah-* (OBJ 2SG, A 3PL)

What we have done in (49) is to add the second person object values to *djirrah-* and *djilah-*. This means that *djirrah-* and *djilah-* are more specific and will therefore fill the 3DU > 2SG and 3PL > 2SG cells.

We now turn to the resolution of CONFLICTS B and C. These are both resolved by the use of rules of referral which are related to the combination of subject and object persons. Recall that we have seen cross-linguistic evidence to suggest that person syncretism in two-place verbs is different from that in intransitives, which strongly suggests that transitivity plays an important role. Because the rules of referral are formulated in terms of the combination of subject and object person they account for the syncretisms observed as resulting in part from transitivity. The rules of referral are as in (50)

- (50) a. $2[n = \alpha] > 1[n = \beta]$ uses the form for $3[n = \alpha] > 1[n = \beta]$
 b. $1[n = \alpha] > 2SG$ uses the form for $3[n = \alpha] > 2SG$
 where n is the number value
 (Evans, Brown and Corbett 2001: 207)

The effect of (50a) is to say that the 2 > 1 paradigm uses the forms of the 3 > 1 paradigm. The second line (50b) states that the 1 > 2SG paradigm uses the forms of the 3 > 2SG paradigm. Hence, in this case the original motivation for the referral is the avoidance of combinations of first and second person objects, the top end of a person hierarchy. The use of variables α , β here is for expository purposes. The point is that the referral does not just involve one cell of the paradigm referring to another. Rather it is a combination of a referral and underspecification. A generalized referral of this kind can then be used to predict whole sub-paradigms.

The referrals in (50) are based on the notion of ‘avoidance’ as elaborated by Heath (1991 and 1998).

The assorted mechanisms ... have in common the fact that they obscure the ‘objective’ relationship between speaker and addressee ... The 1st ↔ 2nd combinations are doubly dangerous because they not only contain the most pragmatically sensitive pronominals, they also combine them into a syntagmatic structure and thereby necessarily focus attention on the speaker–addressee relationship ... The Australian languages ... play down the speaker–addressee relationship by omission, substitution, or skewing of the normal, most transparent, hence also bluntest first and second person morphemes. (Heath 1991: 86)

Dalabon has actually taken this a step further by morphologizing the avoidance strategy, so that there are now no longer unique forms for the $2 > 1$ and $1 > 2$ SG paradigms. This means that one can no longer treat these effects as facts of the ethnography of communication in the Dalabon community, and therefore in the same way as the use of polite *Sie* in German, for example. The politeness rules in German and similar languages can refer directly to the free pronoun (regardless of whether a verb is present in the utterance): the primary effect is that one free pronoun is replaced by another (*du* by *Sie*, and so on), and the effects on the morphology are secondary, simply involving the verb agreeing with the new pronoun in its (etymological) person and number – third plural in German (for interesting complications see Comrie 1975). However, in Dalabon, free pronouns themselves are unaffected by the referrals in (50) which only affect the agreement patterns in verbal prefixation. Consequently, what was once pragmatically systematic in Dalabon has become morphologically systematic syncretism.

Under this analysis, CONFLICT B, in both its specific and its generalized variants, does not arise, because the 1 SG $>$ 2 SG, 1 DU $>$ 2 SG and 1 PL $>$ 2 SG cells are subject to the referral in (50b). The correct forms are selected by rule (50b). As *djah-* is the form used for 3 SG $>$ 2 SG in (29), it will fill the 1 SG $>$ 2 SG cell in (41). As *djirrah-* is the form used for the 3 DU $>$ 2 SG in (29) it will be used for the 1 DU $>$ 2 SG cell in (34), and as *bulah-* is used in the 3 PL $>$ 2 SG cell in (29), it will fill the 1 PL $>$ 2 SG cell in (34). These correct predictions all follow from the single referral in (50b).

Conflict C, in both its specific and generalized variants, is resolved by the rule of referral (50a). The forms used for the 2 DU $>$ 1 SG and 2 PL $>$ 1 SG cells are *burrah-* and *bulah-*, because these are the forms used for the 2 DU $>$ 1 SG and 2 PL $>$ 1 SG cells in (29). The reason why *kah-* is used for the 2 SG $>$ 1 SG cell in (40) is that this is the form used for the 3 SG $>$ 1 SG cell. In the case of referral (50a) it generalizes to all first person object cells where the subject is second person. This can be verified by checking Table 2 where the dual first person object proclitic *njerr* and the plural first person object proclitic *njel* occur in the non-singular first person object portion of the $2 > 1$ paradigm, with in each instance the bound pronominal being the same as for the corresponding $3 > 1$ paradigm (for example, *njerr kah-* is both 2 SG $>$ 1 DU and 3 SG $>$ 1 DU, and *njel burrah-* is both 2 DU $>$ 1 PL and 3 DU $>$ 1 PL).

With the adoption of the referrals in (50), the conflicts resulting from the use of underspecification on its own are resolved. However, the referrals

themselves involve a degree of underspecification, because the number of the argument is underspecified (as indicated by the use of the variables in our informal representation in (50)). Underspecification of number in the presence of person makes cross-linguistic sense, of course, because we know that number marking correlates with person values. The referrals themselves also imply a directionality, in that *burrah-* (A 3 DU), *bulah-* (A 3 PL), *djirrah-* (OBJ 2 SG, A 3 DU) and *djilah-* (OBJ 2 SG, A 3 PL), for example, are now treated as the primary exponents of third person. There is evidence for this. The forms *burrah-* (S 3 DU), and *balah-* (S 3 PL), from which *burrah-* (A 3 DU) and *bulah-* (A 3 PL) are derived by rule (30) are uniquely associated with the third person in the intransitive paradigm. In the singular, *kah-* is also uniquely associated with the third person in the intransitive paradigm, while *djah-*, which marks second person singular object in the transitive paradigm, is uniquely associated with the second person singular subject in the intransitive paradigm. There is no direct evidence that the forms *djirrah-* and *djilah-* are primary exponents of 2 SG $>$ 3 DU and 2 SG $>$ 3 PL respectively, but as they fill these cells, the referral (50b) will make reference to them in any event. While underspecification alone does not work, the generalized (underspecified) referral approach derives the correct results, is favoured by the paradigm-internal evidence and also has a potential original motivation in terms of pragmatically based avoidance based on the top two values of a person hierarchy.

We now turn to a brief account of the implementation of this analysis using DATR.

5.3.4 The verbal hierarchy

This section takes one sample Dalabon verb *nan* 'to see'. Figure 5 illustrates a small part of the lexemic hierarchy in which the verbal lexical items inherit from a node VERB. Like every other verb, verb *nan* 'see' inherits from the node VERB. Its lexical entry is given in (51).

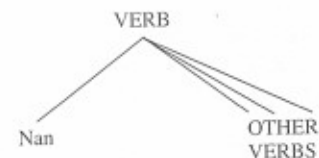


Figure 5. Dalabon verbal hierarchy

- (51) Nan:
 <> == VERB
 <gloss> == see
 <root> == nan.

In (51) the 'empty path' <> is paired with the node VERB, on its right-hand side. This simply means that information which is not specified at the node Nan will be inherited from VERB.⁹ In (52) we give the node VERB.

- (52) VERB:
 <> == undefined
 <mor> == MOR_VERB
 <syn> == SYNTAX
 <syn cat> == verb.

This node serves to bring together the morphological and syntactic specification of verbs in Dalabon. In (52) the path <mor> refers to the path <mor>, and all of its extensions, at MOR_VERB, the node which provides information about the morphology of verbs. It is the information represented at this node which forms the core of our analysis of the Dalabon verbal system and it is to this that we now turn in the next section.

5.3.5 *The shape of the verbal paradigm*

Recall from §5.1.2 that the order of the attributes in our DATR representation is important, because, as a highly constrained kind of underspecification, it determines what can be inferred by default. We saw in (28) in §5.3.1 that there is no person syncretism in the intransitive paradigm. However, there are still constraints on the possible combinations of person and number. Specifically, person imposes certain constraints. First person inclusive is incompatible with singular, because the lowest cardinality associated with it is two, although morphologically the first inclusive dual and the first inclusive plural look like the singular and dual forms respectively of the other persons.

As person determines which numbers are available, we order the person attributes before the number attributes, for a given argument (intransitive subject, transitive subject and object).

⁹ A path cannot be paired with a node without also making reference to another path and its extensions at that node. By convention, where a path is paired with another node and no overt reference is made to a particular path at the node referred to, then the referring path refers to the identical path at the node referenced. In sum, the empty path at the node Nan refers to the empty path, and its extensions, at VERB. Furthermore, the value for any extension of a path which is not already specified at Nan will be found by looking for a matching path at VERB.

- (53) For a given argument of the verb, person attributes are ordered before number attributes.

Given the assumption that underspecification is related to what is cross-linguistically general, (53) entails that there should be some reflex of this cross-linguistically general relationship between person and number in Dalabon. Typically, this would be syncretism of number in the presence of person, but this is not the case for Dalabon. The ordering in (53) is related to our informal description of the referrals in (50). There we used variables to show that the referral does not just involve one cell of the paradigm referring to another, but in fact picks out whole sub-paradigms. The fact that it picks out whole sub-paradigms, of course, shows that the sub-paradigms are defined in terms of person. While avoidance of 1 > 2sg and 2 > 1 is more specific to Dalabon, the involvement of whole sub-paradigms defined by person will follow from the cross-linguistically general principle in (53).

The next step is to determine whether the object attributes should occur before the transitive subject attributes in the paradigm. This question obviously does not arise for intransitive subjects. If we were to base our decision on the ordering of affixes, then the fact that non-singular object clitics appear before the bound prefixes might lead us to claim that object attributes (followed by their person and number attributes) should occur first in the path. We argue, in fact, that this is the case for third and second person object, but not for first person exclusive object.

Only when the object is third or second person may there be clear marking of the person of that object on the bound pronominal prefixes. So the second person singular object is clearly marked on the portmanteau bound pronominal prefix. Furthermore, in the presence of second person singular object marking, the disharmonic transitive subject is syncretic with the dual transitive subject. There is also an example of a bound pronominal prefix clearly marking third person: the third person singular higher animacy object (Table 2). The form *bykah-* is used only when the object is third person singular, not when it is first person exclusive (41), as indicated by the contrast between the 3 > 1 and 3 > 3 paradigms. Hence this form definitely identifies a third person object. In contrast with this there is no bound pronominal prefix form where an object is clearly marked as first person. From this we conclude that objecthood triggers marking of second and third person singular. For the second person singular it also has an effect on the number marking of transitive subjects.

We therefore order attribute combinations of object with second and third person before any attribute combinations of transitive subject (and any associated person and number marking).

There is no unambiguous bound prefixal marking of the first person exclusive object. The bound pronominal system has the same forms for the 3 > 3SG paradigm and the 3 > 1SG paradigm (Generalizations 3–6). In other words, for the bound pronominal paradigm third person transitive subjects condition loss of distinction between first person singular and third person singular objects. This leads us to claim that, in contrast with the other persons, the object attributes for first person exclusive occur last in the path. We shall relate these facts to the referrals in (50). The resulting combination of values in the paradigm is therefore as in (54)

- (54) Second and third person object attributes are ordered before transitive subject attributes¹⁰
 Examples:
 <mor infl o 3rd sg a 2nd sg>
 <mor infl o 3rd du a 2nd sg>
 <mor infl o 2nd pl a 1st_excl sg>

In contrast, first person objects are never distinguished by the bound pronominal morphology and so these are treated as occurring last in the path.

- (55) First person object attributes are ordered after transitive subject attributes
 Examples:
 <mor infl a 2nd sg o 1st_excl sg>
 <mor infl a 3rd sg o 1st_excl du>
 <mor infl a 2nd du o 1st_excl pl>

In (54) and (55) we have examples of the shape of the fully specified paradigm. The attribute ordering reflects what is most likely to be underspecified. We can contrast the avoidance-motivated referral involving the top two positions on the person hierarchy (first and second) with the orderings in (54) and (55), which oppose the lower two positions on the person hierarchy (second and third) with the top position, first person exclusive. The analysis which derives the fully specified combinations, such as those in (54) and (55), can be underspecified in relation to these combinations.

¹⁰ Following Dixon (1994) in our representation we shall use the attributes a, s and o for transitive subject, intransitive subject and object respectively.

5.3.6 Referral of 1 > 2SG to 3 > 2SG

The referral of the 1 > 2SG to the 3 > 2SG combinations illustrates that syncretisms can arise from what is essentially a referral with a degree of underspecification. The underspecified element is the number of the transitive subject. Note that this referral also fits in with the attribute ordering we have determined in (54). As the second singular object attributes are toward the left of the path, they are less likely to be involved in underspecification-based syncretism. This referral is given in (56) below.

- (56) <mor infl o 2nd sg a 1st_excl> == <mor infl o 2nd sg a 3rd>

All extensions of paths for 1 > 2SG can be obtained by referring to the extensions of 1 > 3SG. These extensions will, of course, be those for number of the transitive subject, as this is what is not specified in the referral.

5.3.7 Referral of 2 > 1 to 3 > 1

We claimed in (55) that the first exclusive object attributes should be specified as occurring last in any path. We argued that this had to do with the fact that there is no person-specific bound pronominal marking for the first exclusive object, in contrast with the second and third person at least, where the second singular object is clearly marked and there is marking of the higher animacy third singular object. As this means that first exclusive object marking is an extension of the third person paradigm, the referral of 2 > 1 to 3 > 1 takes the form in (57).

- (57) <mor infl a 2nd> == <mor infl a 3rd>

This states that the second person transitive subject finds its extensions from the third person transitive subject. If third person object attributes were also ordered after transitive subject attributes, this would mean take-over of the 3 > 3 paradigm by the 2 > 3 paradigm. But this is not the case, because from (55) it follows that only first exclusive object attributes can extend a path which starts with transitive subject attributes, such as in (57). Note the high degree of underspecification involved. The referring paradigm of second person transitive subject obtains its transitive subject number, together with the first exclusive object person and number information, from the third person paradigm.

5.3.8 Dalabon: summing up

The referral-based approach allows us to capture the intuition that the syncretism is motivated by the potential inappropriate combination of first

and second person. It is transitivity which conditions the loss of person distinction. It also allows us to treat number marking in the transitive and intransitive paradigms in a similar way. Number distinction is affected by person (first inclusive) in the intransitive paradigm, which makes full person distinctions. The best account requires the use of referrals and underspecification so that whole sub-paradigms may be systematically syncretic. Thus both underspecification and referral are required. It is important to check that our Network Morphology analysis does indeed provide the correct forms for the substantial paradigms of Dalabon. Hence the need for an implementation, a part of which was presented above. The full implementation in Appendix 5 does give exactly the right forms for Dalabon, and only these, and to that extent the Network Morphology account is valid.

Having seen how underspecification can be combined with referral to provide the best account of syncretism in the Dalabon verbal system, we now go on to consider the relationship between position on a morphological class hierarchy and the domains of syncretism. For this we shall use Russian as the main case study, referring to other languages when data from Russian is insufficient.

5.4 Case study 3: The Russian nominal system

The complex inflectional morphology of Russian offers an illuminating case study (58). Its system of inflectional classes provides helpful data for

(58) Main paradigms of Russian nouns

	class I 'law'	class II 'room'	class III 'bone'	class IV 'wine'
NOM SG	zakon	komnata	kost'	vino
ACC SG	zakon	komnatu	kost'	vino
GEN SG	zakona	komnaty	kosti	vina
LOC SG	zakone	komnate	kosti	vine
DAT SG	zakonu	komnate	kosti	vinu
INS SG	zakonom	komnatoj	kost'ju	vinom
NOM PL	zakony	komnaty	kosti	vina
ACC PL	zakony	komnaty	kosti	vina
GEN PL	zakonov	komnat	kostej	vin
LOC PL	zakonax	komnatax	kostjax	vinax
DAT PL	zakonam	komnatam	kostjam	vinam
INS PL	zakonami	komnatami	kostjami	vinami

discussing the notions of domain and regularity. Russian has four basic noun paradigms, as given in (58).

In a traditional account, these paradigms might be treated as monolithic units. In a Network Morphology account, we would analyse the oblique plural forms as being shared across the paradigms; thus the dative plural of a noun consists of its stem plus *-am*, irrespective of inflectional class. We would also treat the nominative plural as being stem plus *-i* (or its orthographic variant *-y*), with this being overridden just for inflectional class IV. At a lower level, we would capture the shared forms of inflectional classes I and IV. That is, we would have a hierarchy of defaults, the highest applying very generally, in fact applying to more than just nouns, the lowest having smaller domains, and at the bottom of the hierarchy would be lexical items which must contain some idiosyncratic information. A possible structure is given in Figure 6. In addition to the four lexical items given, there are thousands of other nouns which inherit from the four inflectional class nodes. For the detail see Corbett and Fraser (1993). The implementation on which this case study is based is from Brown (1998b), and a simplified version of this is given in Appendix 6.¹¹ The relevance of this approach is that it suggests a range of possible domains for syncretism.

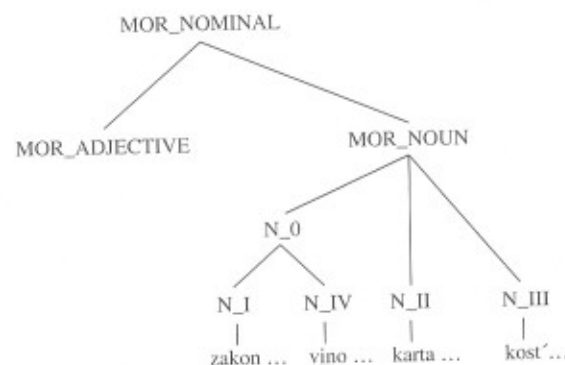


Figure 6. An inheritance structure for Russian nominals

¹¹ The fragment *rusnoms.dtr* at the Sussex DATR archive <http://www.cogs.sussex.ac.uk/lab/nlp/datr/datr.html> precedes Brown (1998b) but still involves separate hierarchies. Corbett and Fraser (1993) gave a single hierarchy. Later papers use a network of hierarchies; in particular a syntactic hierarchy is added, which is concerned with the syntactic category of items; by default, items which are, for example, syntactic nouns will inherit information from the noun section of the morphological hierarchy, but this is not always so.

The higher up the hierarchy that a syncretism is stated, the larger the domain of syncretism. And the larger the domain, the more systematic the syncretism.

5.4.1 *Domains of syncretism*

Given that we find numerous instances of syncretism, what are the possible domains? We first separate off phonology and then move to the central potential domains.

5.4.1.1 Phonologically determined 'syncretism'

Certain syncretisms can be excluded on the grounds that they result solely from a phonological rule; for example, the collapse of genitive singular with nominative/accusative singular in some stem-stressed nouns in Russian, discussed in Chapter 1: §1.5, which is due to the reduction of unstressed vowels. Such instances can be straightforwardly accounted for by the phonology and therefore excluded from treatment within morphology proper.

5.4.1.2 Lexically determined syncretism

The least systematic type of syncretism (that involving the smallest domain) would be a syncretism for a single lexical item. We do not find such an instance in Russian. To illustrate what this would look like, consider the paradigm of Italian *essere* 'be' in (59).

(59) Italian *essere* 'be' in comparison with a regular verb

		'be'		'speak'	
		singular	plural	singular	plural
1		sono	siamo	1	parlo
2		sei	siete	2	parli
3		è	sono	3	parla
					parlano

According to Davide Ricca (personal communication) the verb *essere* 'be', which has first person singular and third person plural both as *sono*, is the only verb in the language to have this identity. This then is a case of lexically determined syncretism, which is equivalent to Coleman's (1991) first degree syncretism. Other languages show further unique syncretisms, though these are rare.

In principle the situation is the same, whether it affects a single item or several items in an unmotivated way: that is, the syncretism would need to be specified in the lexical entry for each of them. Here is a further example of lexically determined syncretism. Numerals in Russian vary in the number of case distinctions made, but the normal minimal paradigm distinguishes three forms. Consider these numerals against that general background:

While syncretism of nominative and accusative is widespread in Russian numerals, syncretism of all the remaining case forms is unusual, and is the only morphological feature these items have in common. This combination of syncretisms including the instrumental must be lexically specified for these three items only.¹²

(60) Lexically determined syncretism in Russian numerals

	'forty'	'hundred'	'ninety'
NOM	sorok	sto	devjanosto
ACC	sorok	sto	devjanosto
GEN	soroka	sta	devjanosta
DAT	soroka	sta	devjanosta
LOC	soroka	sta	devjanosta
INS	soroka	sta	devjanosta

5.4.1.3 Morphologically determined syncretism

In this part of our case study of Russian we illustrate how morphological regularity is itself a matter of degree. Referring back to the Russian paradigms in (58) we now consider morphological domains, starting from the smallest. The smallest morphological domain above that which is lexically determined is syncretism within a single inflectional class. We find this in the class we have labelled class III in (58), where there is, among other things, identity of genitive singular and locative singular, repeated here in (61).

¹² *Sorok* 'forty' and *sto* 'hundred' are the clear cases; because of vowel reduction all the forms of *devjanosto* 'ninety' are pronounced identically, thus there is an additional phonologically induced syncretism of the nominative and accusative with all the other cases, giving only one phonological form for this item. Note also that GEN/DAT/LOC/INS syncretism is also characteristic of the singular feminine adjectives; see (67) below.

We could reflect this in the lexical entry. However, we would be suggesting, counterfactually, that the identity is an idiosyncratic fact about the particular noun *kost'* 'bone', which is not shared with other items, in other words that the syncretism is lexically determined. In fact, the syncretism of

(61) Syncretism within a single inflectional class in Russian

	class III 'bone'
NOM SG	<i>kost'</i>
ACC SG	<i>kost'</i>
GEN SG	<i>kosti</i>
DAT SG	<i>kosti</i>
LOC SG	<i>kosti</i>
INS SG	<i>kost'ju</i>

locative and genitive singular is one of the characteristics shared by all members of the inflectional class (Coleman's 1991 second degree syncretism). We should therefore state this syncretism at a higher point in the inheritance hierarchy in Figure 6, at the node *N_III*, from which *kost'* inherits (as do over 4,000 other nouns). This syncretism is evidently more systematic than any lexically stipulated syncretism. It is, however, less systematic than instances we shall come to shortly.

There is also a good reason to treat the syncretism in (61) as directional. This is represented with the DATR equation in (62). Again, we have omitted some information, as indicated by the ellipses.

(62) *N_III*:
 <> == MOR_NOUN
 <mor sg loc> == "<mor sg gen>"
 ...

The equation in (62) is a referral which states that for class III the form of the locative singular is determined by the form of the genitive singular. The reason for this is that the form of the genitive singular *-y* in class II, although orthographically different in (58), can be treated as phonologically identical with the exponent *-i* of class III, suggesting that the stem *+i* combination is primarily the exponent of genitive singular, and secondarily in class III the exponent of locative singular. This analysis is further justified when we consider syncretism which ranges over more than one inflectional class.

In Russian, as can be seen in (58), dative and locative singular are identical in two inflectional classes (II and III). We repeat the relevant paradigms in (63).

(63) Syncretism within a subset of the inflectional classes of a word class

	II 'map'	III 'bone'
NOM SG	<i>karta</i>	<i>kost'</i>
ACC SG	<i>kartu</i>	<i>kost'</i>
GEN SG	<i>karty</i>	<i>kosti</i>
DAT SG	<i>karte</i>	<i>kosti</i>
LOC SG	<i>karte</i>	<i>kosti</i>
INS SG	<i>kartoj</i>	<i>kost'ju</i>

While the case forms are identical within the inflectional classes, the inflections involved differ between inflectional classes. This example is more systematic than that involving just class III. It is more systematic in two ways. First, many more nouns are involved, all of those in classes II and III. We state the syncretism at the *MOR_NOUN* node (and then override it for the remaining nouns which inherit from *N_O* in Figure 6, see Brown (1998b: 257)). Second, the syncretism holds true for two quite different morphological realizations: those nouns like *karte* 'map.DAT SG/LOC SG' and those nouns like *kosti* 'bone.DAT SG/LOC SG', as shown in (63).

If we refer back to the example paradigms in (58), we see that the locative singular realization combining a stem with the ending *-e* is the default for nouns as a whole, because it is found in classes I, II and IV. Furthermore, in classes I and IV, the combination of a stem with the ending *-e* is reserved solely for the locative singular. This is repeated in (64).

(64) The ending *-e* as primary exponent of locative singular

	I 'law'	IV 'wine'
NOM SG	<i>zakon</i>	<i>vino</i>
ACC SG	<i>zakon</i>	<i>vino</i>
GEN SG	<i>zakona</i>	<i>vina</i>
DAT SG	<i>zakonu</i>	<i>vinu</i>
LOC SG	<i>zakone</i>	<i>vine</i>
INS SG	<i>zakonom</i>	<i>vinom</i>

Examination of the paradigms in (63) and (64) suggest that the combination of stem and ending *-e* should primarily be associated with locative singular as a default for nouns.

The DAT/LOC syncretism is therefore a referral of dative singular to locative singular. This can be expressed as in (65).

```
(65) MOR_NOUN:
      <> == MOR_NOMINAL
      <mor sg dat> == "<mor sg loc>"
      <mor sg loc> == "<stem sg>" e "<stress sg>"
      ...
```

Class II, which includes nouns such as *karta* 'map', inherits both equations and therefore combines the syncretic pattern with the default ending *-e*. The intermediate class *N_O*, from which class I and class IV inherit, overrides the referral by stipulating that the realization of dative singular is the stem plus the ending *-u*. Class III, on the other hand, overrides the default realization of the locative singular, while still inheriting the default referral of dative singular to locative singular. When the referral in (65) is combined with the referral specific to class III in (62) we obtain a chain in which the dative singular is based on the locative singular, and the locative singular is based on the genitive singular, thereby yielding the collapse of three case distinctions in class III. Analytically this is important, because the identity of genitive and dative in class III is the product of combining the noun default referral for DAT/LOC with the class III referral for LOC/GEN.

With the original Jakobsonian features the DAT/LOC syncretism can be captured only by leaving the realization specified as +peripheral. However, this would also include the instrumental case. In contrast, it is easier to pick out the LOC/GEN syncretism. The GEN/DAT syncretism in class III is problematic for the Jakobsonian feature system, as these cases do not share any values under the unmodified version. It is possible to express some of these generalizations using a modified variant of the Jakobsonian approach (Chapter 3: §3.1.1). The problem is that a decision must be made about which generalizations are to be captured, and this involves precluding other phenomena. For instance, in Müller's (forthcoming) approach to syncretism and inflectional allomorphy in Russian, there are two separate vocabulary items, /e/4 and /e/5, the former used for the dative/locative of class II and the latter used for the locative of classes I and IV. The reason for this is that, even though dative and locative form a natural class [-subj, +obl], Müller also attempts to capture sharing between inflectional

classes by using variables, and the variables for classes I and IV exclude those for class II. In fact, scaling up underspecification-based approaches is often a problem when taken in the context of other features. Breaking the features up even further to make underspecification work brings with it the potential for massively increasing the number of elsewhere, which the system underutilizes in any event. When taken in isolation this problem for theories which manipulate the feature values is not as great as when number and case, or lexical information about inflection class, are actually combined, thereby yielding a massive number of potential combinations which never appear. In contrast, the analysis of the noun syncretisms based on atomic values and referrals actually captures the licensing effect that one syncretism may have on others.

Having seen an example of a syncretism which covers more than one morphological class within a word class, we now turn to a more general domain, namely to potential examples of syncretism throughout a word class. For a clear case we turn to Russian's South Slavonic relative, Serbo-Croat. Here the accusative plural and genitive plural are syncretic for the personal pronouns, but not for nouns and adjectives. In (66) we illustrate this syncretism using the first and third person pronouns (non-clitic).

(66) Syncretism throughout one word class in Serbo-Croat

	1PL	3PL
NOM	mi	oni (M), ona (N), one (F)
ACC	nas	njih
GEN	nas	njih
DAT	nama	njima
LOC	nama	njima
INS	nama	njima

The other pronouns behave in the same way with regard to the PL ACC/PL GEN syncretism. Thus we have an instance which involves all members of the word class. While the word class involved in Serbo-Croat is small with respect to the number of lexical items involved, we have already seen an instance where the lexical class was much larger, namely Finnish (example (13) in Chapter 3: §3.1.2), where syncretism affects all nouns, but not pronouns.

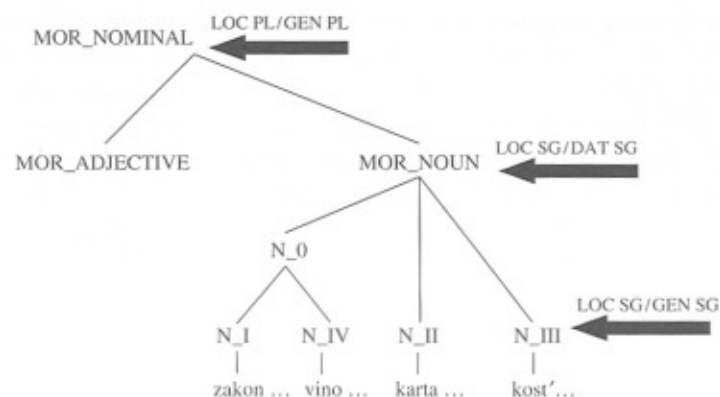


Figure 7. *The location of syncretism patterns in the Russian inheritance hierarchy*

We also find syncretisms which generalize across more than one word class. In Russian all adjectives and pronouns have the genitive plural syncretic with the locative plural. This does not extend to nouns, where the inflected forms differ, as can be seen in (58).

Finally, it is possible to find examples of syncretism across all potentially relevant word classes. In Slovene, the dative dual and instrumental dual are syncretic for nouns, adjective and pronouns (Priestly 1993: 399). That is to say, these forms are syncretic for anything which can mark them. No lexical item of Slovene has a dative dual which is distinct from the instrumental dual. The type of inheritance hierarchy illustrated in Figure 6 suggested possible morphological domains for syncretism, and we have found instances of all of them. In Figure 7 we show where the generalizations are located on the inheritance hierarchy for Russian. The syncretism of locative plural and genitive plural is a property of nominals, because it occurs for pronouns as well as adjectives, but it is overridden by nouns. The locative singular and dative singular syncretism is a default for nouns, but overridden at the node N₀, and the syncretism of the locative singular and genitive singular is a default for class N_{III} only and is specified there. The locations in the hierarchy represent different degrees of regularity. The higher up the hierarchy a syncretism is stated, the greater the number of items that are likely to be affected. The second aspect to regularity, namely that with inflectional classes a syncretism may generalize over different realizations, was illustrated in (63) above, and we shall discuss a more dramatic instance of it in §5.4.2. Before that we must

consider what would be a highly systematic type of syncretism, but one for which we have found no convincing evidence.

5.4.1.4 Candidates for syntactically determined syncretism

There are various types of situation which at first appear to be syntactic domains for syncretism. On closer examination they seem to us to be better analysed in other ways.

In the earlier discussion we talked of 'noun', 'adjective' and so on, which suggests that syntactic categories can provide the domain for syncretism. However, the hierarchy given in Figures 6 and 7 simplifies the position. We need a separate hierarchy for syntactic categories, though by default items in a given category inherit their morphological information from a node dominated by a morphological category corresponding to this syntactic category (nouns from MOR_NOUN, adjectives from MOR_ADJECTIVE, and so on). The question then arises as to what happens when syntactic and morphological categories fail to match. Which provides the domain for syncretism? Slavonic provides helpful data here. Consider Russian items like *stolovaja* 'dining room'. Syntactically these behave like nouns. However, they behave morphologically as adjectives (67).

(67) Russian nouns and adjectives

	'map'	'dining room'	'new' (F.SG)
<i>syntactically:</i>	noun	noun	adjective
<i>morphologically:</i>	noun	adjective	adjective
NOM SG	karta	stolovaja	novaja
ACC SG	kartu	stolovuju	novuju
GEN SG	karti	stolovoj	novoj
DAT SG	karte	stolovoj	novoj
LOC SG	karte	stolovoj	novoj
INS SG	kartoj	stolovoj	novoj

The syncretisms found with morphological nouns and morphological adjectives are different in Russian. *Stolovaja* 'dining room' and similar items have the syncretism of morphological adjectives. This shows that it is the *morphological* category which provides a domain for syncretism. To demonstrate that a syntactic category is a possible domain, we would need to find the converse situation to that found in Russian, that is a situation in

which morphological category and syntactic category differed for certain items, and where these items took the syncretisms appropriate for the syntactic category but not for the morphological category.¹³

The other area where we might consider syntactic influence on syncretism is with features which may have a role both in morphology and in syntax. (They are to be distinguished from purely morphological features, such as features indicating inflectional class.) Typical examples are number and gender, which are relevant to syntax, notably in agreement, and also to morphology. To see the role of such features we shall consider the interesting issue of animacy, discussed above in Chapter 3: §3.1.2 and Chapter 4: §4.3.4 (also see Perlmutter and Orešnik 1973, Huntley 1980, Corbett and Fraser 1993).

Table 3 includes two examples for each of the four main inflectional classes (see (58) above); each of the eight examples is representative of a group of nouns. If we look at any of the examples individually, we find an instance of syncretism. Take just the first example, *student*. There is syncretism of accusative and genitive singular (conditioned by animacy). More generally, the singular accusative depends on the gender and animacy of the item in question in Table 3, and the plural accusative on the animacy of the item in question. Evaluation of this information allows us to state that the singular accusative of masculine animates is the same as the singular genitive, and the plural accusative of animates is the same as the genitive.

The different inflectional paradigms in Table 3 share the same *pattern* of identity, even though the particular inflections differ. It would clearly be inadequate to state the identity of forms separately for each inflectional class; that would imply that the cases involved could equally well differ from inflectional class to inflectional class. This regularity is captured in the Network Morphology account by a statement high up the inheritance

¹³ Andrew Carstairs-McCarthy (personal communication) points out a possible example of syntactically conditioned syncretism in Latin. Some third declension adjectives have syncretism of the dative and ablative singular (with the ending *-ī*) when used adjectivally, but when used substantivally, these forms are distinct (dative *-ī* versus ablative *-e*), e.g. *ā sapienti viro* 'by a wise man', *ā sapiente* 'by a philosopher' (Kennedy 1955: 40). However, this alternation also obtains *within* nouns of the third declension. The third declension is divided into two subtypes, the *i*-stems and the consonant stems: *i*-stems have the syncretic dative/ablative singular *-ī*, while consonant stems have a distinct dative singular *-ī* versus ablative singular *-e*. Thus, the choice of forms is not necessarily correlated with syntax, and the alternation between syncretic and non-syncretic patterns in the third declension adjectives might better be described as an alternation between *i*-stem and consonant-stem declension patterns.

Table 3 *The morphological effect of animacy in Russian*

	'student' animate	'law' inanimate	'teacher(F)' animate	'map' inanimate	'mouse' animate	'bone' inanimate	'monster' animate	'wine' inanimate
NOM SG	student	zakon	učitel'nica	karta	miš	kost'	čudovišče	vin
ACC SG	studenta	zakon	učitel'nicu	kartu	miš	kost'	čudovišče	vin
GEN SG	studenta	zakona	učitel'nicy	karty	miši	kosti	čudovišča	vina
NOM PL	studenty	zakony	učitel'nicy	karty	miši	kosti	čudovišča	vina
ACC PL	studentov	zakony	učitel'nic	karty	mišej	kosti	čudovišč	vina
GEN PL	studentov	zakonov	učitel'nic	kart	mišej	kostej	čudovišč	vin

tree. Since animacy also affects the agreement of adjectives, and the form of pronouns, the logical place for it would be at the top of the hierarchy in Figure 7, namely at the MOR_NOMINAL node.

There are two questions here, the first is the type of feature we are dealing with, and the second, our main concern, is the domain of syncretism. Animacy appears to be a semantic feature, in that the nouns involved denote entities which live and move (thus insects are animate but plants are not). The match with this semantic definition is close in Russian (less so in some other Slavonic languages). There are some interesting borderline cases, for instance *pokojnik* ‘the deceased’ is grammatically animate. Such instances are animate for agreement purposes and for the morphological matter of syncretism. But we should also capture the fact that the personal pronouns regularly have accusative/genitive syncretism, whether or not they are referring to an animate entity. This may be captured by stipulating that they are grammatically animate. Our conclusion is that the animacy feature is a morphosyntactic one (albeit one with strong semantic motivation in Russian, and with lesser semantic motivation in some other Slavonic languages). In any case, while animacy is a major determining feature for accusative/genitive syncretism, the syncretism depends on the interaction of animacy with number and gender (see Fraser and Corbett 1995 for more details). Moreover, it is overridden by morphological considerations (inflectional class II nouns have accusative singular in *-u* irrespective of animacy). In our previous examples the domain given uniquely specified the syncretism in question. The animacy feature does not do this. What then is the domain of the syncretism we have just examined? We shall see in the next section.

5.4.2 *Orthogonal specification of syncretism*

The clearest domains we have encountered were well described in terms of an inheritance hierarchy. However, the interesting data on animacy prove significantly different. The point is that generalizations about syncretism determined by animacy can be stated high up at the level of MOR_NOMINAL in the hierarchy in Figure 7. They apply to different paradigms, giving identities of pattern rather than any phonological identity. It is not the case that accusative/genitive syncretism is found with all items in the nominal domain, rather only those that are animate (whether as a result of their semantics, or being specified as animate, or being marked as animate as a consequence of agreement). Thus the specification is orthogonal to the specification for inflectional class.

This captures the common-sense view that we do not wish to claim there are eight inflectional classes for nouns in Russian, rather that there are four main classes, with animacy affecting each of them.

The natural way of thinking of domains is in hierarchical terms, and this was our approach. However, syncretism may require specification which is orthogonal to the morphological hierarchy. How then does such orthogonal specification of syncretism differ from lexical specification? Lexical specification means that the individual items must each be specified, in other words that they are exceptional in this regard. Orthogonal specification can be regular (animacy syncretism in Russian shows a very high degree of regularity). The distinguishing point is that it depends on a feature which is otherwise irrelevant to inflection, because it affects only the accusative case, and which cross-cuts the features which determine inflection. We can specify a domain *within* which it operates, but not a domain where it uniquely determines a syncretism.

Animacy in Russian shows that regularity of syncretism is quite a subtle notion. On the one hand, syncretism based on animacy is systematic, because its evaluation potentially determines the accusative plural of every lexical item which is nominal, and therefore it is stated high up in the hierarchy. Furthermore, it is realized in several different ways which have no phonology in common. On the other hand, there are many nouns, for which in the singular it has no effect (as already mentioned, inflectional class II nouns have accusative singular in *-u* irrespective of animacy).

5.5 Conclusion

We have examined three case studies of unrelated languages, illustrating different aspects of syncretism. The Dhaasanac study required the generalization of the A/B pattern across different tenses. Furthermore, this was a pattern which, while systematic, could not be accounted for using morphosyntactic features alone. By separating the realization of the forms from the morphosyntactic feature specification, Network Morphology makes it possible to account for such systematicity. As it places constraints on feature specification, Network Morphology is still able to relate this language-specific morphological systematicity to the general patterns: the choice of A/B patterns is determined by tense, aspect, mood and negation. In our case study of Dalabon we showed that what was originally a pragmatically determined avoidance of particular forms – namely those involving combinations of first person and second person in the transitive

paradigm – had become morphologized. Underspecification cannot deal with this syncretism on its own. The Network Morphology construct of generalized referral involves a degree of underspecification and is therefore able to refer one set of cells in the paradigm to another set of cells. This means that it is not just stating the facts directly, but predicting whole sets of forms. Finally, our case study of Russian showed that degrees of morphological regularity can be defined by locating syncretisms at different places in an inheritance hierarchy. Each of the case studies in this chapter has an implementation associated with it. The fragment for Dhaasanac in Appendix 4 generates the positive perfect and imperfect paradigms, as well as the positive imperative singular, for three example verbs, two of the coronal type and one of the non-coronal type. The Dalabon fragment in Appendix 5 covers 117 forms (102 transitive and 15 intransitive) derived for the verb 'to see', and the Russian fragment in Appendix 6 covers a significant proportion of the nominal lexicon. Our approach therefore has justification on two grounds. We have put forward linguistic arguments which favour these approaches, and we have implemented our analyses to demonstrate that linguistic arguments are indeed tenable.