Reprinted from

Lingua

Lingua 103 (1997) 25–57

Defaults in Arapesh[☆]

Norman M. Fraser, Greville G. Corbett

Department of Linguistic and International Studies, University of Surrey, Guildford, Surrey GU2 5XH, UK

Received 4 April 1995; revised version 29 May 1997

N·H
ELSEVIER

ELSEVIER

Lingua 103 (1997) 25–57

Lingua

Defaults in Arapesh☆

Norman M. Fraser, Greville G. Corbett

Department of Linguistic and International Studies, University of Surrey, Guildford, Surrey GU2 5XH, UK

Received 4 April 1995; revised version 29 May 1997

Abstract

Network Morphology is a formally explicit approach to morphology which distributes information across a network in which generalizations can be optimally expressed. Generalizations become available in specific cases by the operation of default inheritance. In this paper we explore the notion of 'default' in morphology by means of a Network Morphology analysis of the noun classes and genders of Arapesh – a language which relies on a sophisticated understanding of defaults for a satisfactory treatment (Aronoff, 1992). Our work lends support to Aronoff's account of the Arapesh data. It also reveals a confusion in use of the term 'default' by linguists. In one usage of the term, the ('normal case') default is that which applies in the absence of blocking information; in the other, the ('exceptional case') default is that which applies when some exceptional factors prevent normal processes from applying and necessitate the adoption of some 'last resort' solution. Under one reading the default equates with the unmarked case; in the other, it is the marked case which is picked out by the same term.

1. Introduction

In previous papers we have introduced Network Morphology, a theoretical framework in which default inheritance plays a central role. We have shown how the complex nominal inflection of Russian can be analysed in this framework (Corbett and Fraser, 1993) and have presented a formal account of default assignments of gender and animacy, and default assignments to morphological class for Russian nouns (Fraser and Corbett, 1995).

☆ Earlier versions of this paper were read at the Autumn Meeting of the Linguistics Association of Great Britain, Bangor, 13–15 September 1993, and at the Second ESRC Research Seminar on 'Frontiers of Research in Morphology', London, 3 March 1995. This research was supported by the Economic and Social Research Council (grants R000233633 and R000236063) and the Leverhulme Trust (grant F.242M). We are grateful to them for their support. We also wish to thank Dunstan Brown, Gerald Gazdar and an anonymous reviewer for helpful comments on a draft of this paper.

PII S0024-3841(97)00012-0

Our work has uncovered a surprising amount of flexibility in use of the term 'default' in earlier morphological accounts, including our own. In this paper we explore the notion of 'default' in morphology by means of a Network Morphology analysis of the noun classes and genders of Arapesh – a language which requires some understanding of defaults for a satisfactory treatment (Aronoff, 1992).

We begin by presenting the central concepts of Network Morphology (Section 2), and particularly the place of default inheritance in that framework. We also introduce DATR, the formal language used for encoding Network Morphology theories. In Section 3, we describe some features of the noun morphology of Arapesh. We then offer a Network Morphology analysis of the Arapesh data in Section 4 (morphological class assignment) and in Section 5 (gender assignment). To conclude (Section 6), we identify two distinct ways in which the term 'default' has been used in our account, and argue that the term is insufficiently precise to avoid confusion, unless it is further specified.

2. Network Morphology

Network Morphology is an approach to morphology which distributes information across a network in which generalizations can be optimally expressed. Generalizations become available in specific cases by the operation of default inheritance.

Network Morphology theories are expressed in a formal representation language called DATR which was developed by Roger Evans and Gerald Gazdar at the University of Sussex. DATR is a particularly useful formalism for developing explicit accounts of complex linguistic data because it is formally well-defined (Evans and Gazdar, 1989a) and it allows for the construction of largely declarative accounts which rely on a limited set of basic operations, of which default inheritance is one (Evans and Gazdar, 1989b).[1] Added to the formal rigour and rich expressiveness of the DATR language is a third feature whose importance should not be overlooked by theoretical linguists: computer interpreters (and compilers) exist which are capable of taking a linguistic theory expressed in DATR as input and automatically generating as output all the forms which the theory allows. To work with such a tool is to do no more than take seriously some of the foundational precepts of generative grammar. Whilst many linguists display a remarkable capacity for memorizing and recalling facts about language and languages, we are sceptical about the abilities of any human being to follow through all the ramifications of every minor change to

[1] DATR has been one of the major sources of inspiration in the development of the framework of Network Morphology. We gladly acknowledge our debt to Roger Evans and Gerald Gazdar for this. Helpful introductions to DATR for linguists can be found in Evans and Gazdar (1996) and Gazdar (forthcoming). Prior to our contact with DATR, Fraser's (1985) work on inheritance in Word Grammar (Hudson, 1984) was formative in our thinking. Though it emerged independently, our work on Network Morphology has many features in common with Gibbon's ILEX (integrated lexicon) programme. Gibbon envisages the ILEX framework as "a set of linguistic constraints on the form of possible DATR representations" (Gibbon, 1992: 47).

today's complex linguistic theories.[2] As linguists who have tried both the 'deep thought' and the 'computer assisted' methods of working, we are convinced of the advantages of the latter. Having said this, it is important to stress that the venture we are engaged upon is still unambiguously *theoretical* linguistics. Our primary concern is with linguistic issues of observation, description and explanation rather than computational issues such as algorithmic efficiency.

As a first example, consider the morphological encoding of plural number in English. The vast majority of count nouns form their plural by the addition of an *-s* suffix (which is realised as /-ɪz/, /-z/ or /-s/ according to regular phonological rules). A small number of lexical items form their plural by other means. A particularly parsimonious way of expressing this is to state the general plural formation rule once and to allow it to apply to all lexical items which do not include an exceptional plural form. The general rule is the **default**; it applies in every case where it is not **blocked** (or **overridden**) by a specified exception.

The facts of English number morphology can be expressed simply in DATR as shown in (1) (only a representative sample of lexical entries is shown and only facts relevant to the immediate discussion are included; for simplicity of presentation orthographic representations are used):

```
(1) COUNT_NOUN:
      <> == NOUN
      <plural> == "<singular>" s.

    Dog:
      <> == COUNT_NOUN
      <singular> == dog.

    Cat:
      <> == COUNT_NOUN
      <singular> == cat.

    Horse:
      <> == COUNT_NOUN
      <singular> == horse.

    Ox:
      <> == COUNT_NOUN
      <singular> == ox
      <plural> == oxen.

    Sheep:
    <> == COUNT_NOUN
      <singular> == sheep
      <plural> == <singular>.
```

2 The horrifying (and all too familiar) experience of discovering an egregious and disastrous error in a linguistic theory which has taken years to develop is captured with chilling precision in Pullum's satirical essay 'The incident of the node vortex problem' (Pullum, 1989).

A DATR theory (such as this one) consists of a collection of **facts**. A typical fact has three components: a node, a path and a value. A **node** is a symbol which appears to the left of a colon (':'). Conceptually, it is a location in a network structure at which one or more facts may be stored. Where several facts are stored at the same location, repeated node labels and colons may be omitted so long as the list of facts stored at that node is written as a contiguous block terminated by a full stop. Thus, the facts stored at the Ox node in (1) could equally have been represented as in (2).

(2) Ox: <> == COUNT_NOUN
Ox: <singular> == ox
Ox: <plural> == oxen.

All node names begin with an upper case character. By convention, nodes corresponding to lexical entries have an initial capital followed by lower case letters. All other nodes are fully capitalized. Node names must be atomic, that is, they must consist of a sequence of characters with no intervening spaces. The underscore character ('_') can be used to improve readability while retaining atomicity.

A **path** is that part of a fact which appears immediately to the left of the double equals ('==') symbol. It consists of zero or more atoms enclosed in angle brackets. The maximum number of atoms appearing in the paths shown in (1) is one (for example <plural>). However, any number of atoms may appear. Thus, <accusative plural> would be a well-formed path in this notion. Paths in DATR can be thought of as increasing in specificity from left to right. Thus, <accusative> is a description which includes accusative nouns irrespective of number, whereas <accusative plural> applies only to those accusatives bearing plural number. In this way, the length of a path may be set to achieve the desired degree of specificity of description. An empty path ('< >') is maximally underspecified and so describes everything capable of description in DATR. The function of the empty path will be explained shortly.

Values appear to the right of the double equals symbol. Values may be referenced directly or indirectly. A direct reference to a value is simply a statement of the value. It may consist of an atomic symbol (such as dog or cat) or a list of values (for example, the component formatives of *friendliness* could be presented as friend ly ness, where friend, ly and ness are distinct atoms).

A value may be referenced indirectly by listing in its place a path at which the desired value may be found. For example, the syncretism between past tense and past participle in English regular verbs could be expressed as shown in (3).

(3) VERB:
<past_participle> == <past_tense>.

This states that the value of the <past_participle> path is whatever the value stored at the <past_tense> path happens to be. Because no node name is specified for the <past_tense> path, it is taken to be available at the VERB node (i.e. it is taken to mean VERB: <past_participle>). It is possible to reference a value stored at another node by prefixing that node name to the path name. For example, the identity of inflection in Russian between first declension nominative singular nouns (N_I) and third declension nominative singular nouns (N_III) could be presented as shown in (4).

```
(4) N_III:
      <mor nom sg> == N_I: <mor nom sg>.
```

This captures the facts correctly, but is capable of more efficient representation, as depicted in (5) (this is the analysis offered in Corbett and Fraser, 1993: 137).

```
(5) N_III:
      <mor nom sg> == N_I.
```

By convention, when a node name appears to the right of a double equals symbol without a following explicitly referenced path name, the name of the path which appears to the left of the double equals symbol is implicitly understood to be referenced at the node to the right of the double equals symbol.

Consider again the fragment shown in (1) and, in particular, the first two nodes in the fragment. These are reproduced here as (6).

```
(6) COUNT_NOUN:
      < > == NOUN
      <plural> == "<singular>" s.

    Dog:
      <> == COUNT_NOUN
      <singular> == dog.
```

The first fact listed at the Dog node consists of the node name (Dog), the empty path ('<>') and another node name (COUNT_NOUN). From what we have just seen we know that Dog: <> == COUNT_NOUN is exactly equivalent to Dog: < > == COUNT_NOUN: <>. Earlier we noted that paths are descriptions which increase in specificity from left to right, so the empty path is maximally underspecified. Put simply, this fact reads as follows: the value of anything not described at the Dog node is stored at the COUNT_NOUN node. Thus, if we were to ask the question 'What is the singular of *dog*?' this would amount to searching in the network for a fact of the form Dog: <singular> == ?, where '?' stands in place of the value for which we are searching. A fact matching this pattern is found immediately at the Dog node, namely Dog: <singular> == dog. We need look no further; more than this, we must look no further since the first matching fact we encounter blocks the acquisition of any further facts.

Now suppose we want to know the plural of *dog*. This amounts to searching for a fact of the form Dog: <plural> == ?. No <plural> path is available at the Dog node, but the catch-all empty path serves as a minimal description of <plural>. This makes the default condition available: the value of anything not described at the Dog node is stored at the COUNT_NOUN node. Thus, the value of Dog: <plural> is referenced to COUNT_NOUN: <plural>. The <plural> path is specifically mentioned in a fact stored at COUNT_NOUN, where its value is given as "<singular>" s. This is a list-structured value consisting of an indirectly referenced value "<singular>" followed by a directly referenced atomic value (s).

Quoted paths such as "<singular>" serve a very useful purpose: they provide a means of referencing the value of the path *at the node where the query began*. In this case the query began at the Dog node. We have already learned that the value of <singular> at Dog is dog, so we may conclude that the value of <plural> at Dog is dog s.

The <plural> == "<singular>" s fact stored at the COUNT_NOUN node serves as a **default** which may be **inherited** by lexical entries, except in those cases where the default is overridden (that is, pre-empted) by a more specific, and therefore exceptional, fact. In example (1), default inheritance makes the regular plural schema available at Dog, Cat and Horse. No attempt is ever made to inherit the regular plural form for Ox because it is overridden by the exceptional plural form oxen which is listed at the Ox node. Likewise, the unattested regular plural **sheeps* is overridden at the Sheep node by the exceptional singular-plural syncretism described by the fact <plural> == <singular>.

Examples dealing with data as simple as English plural marking serve to introduce the basic terminology and notation of our framework; they do little to demonstrate how powerful and suggestive it can be. This has been shown in a series of papers in which Network Morphology has formed the basis of unifying accounts of some of the more complex phenomena relating to the nominal morphology of Russian (Corbett and Fraser, 1993; Brown and Hippisley, 1994; Fraser and Corbett, 1995; Brown et al., 1996).

Network Morphology furnishes a descriptive framework which allows complex morphological systems to be described with considerable parsimony. If this were its only asset it would be no more than an efficient descriptive system. However, Network Morphology also embodies a set of restrictive statements which place constraints on the forms which linguistic theories may take. One of the ongoing aims of the Network Morphology programme is to enrich this stock of linguistically-motivated constraining principles and build these (along with formal reconstructions of previously established principles) into a coherent formally explicit theory of morphology.

For example, Network Morphology is typologically informed. One of the principles we assume is that for all languages which mark gender, the gender of the majority of nouns can be predicted from information required independently in the lexicon (Corbett, 1991: 68). On the basis of those languages which have been studied in depth, we take 'the majority' to mean at least 85% of nouns. In addition, we claim that all gender assignment systems have a semantic core. The content of this claim is that in cases when there is a clash between a gender assignment rule based on semantics and one or more gender assignments based on other features (such as phonology or morphology), the semantic assignment takes priority. We have built these typological principles into our analysis of Russian (Fraser and Corbett, 1995).

In the next section we describe the gender system of Arapesh, a language rich in genders, and incorporating a number of features which make it ideal as a testbed for Network Morphology.

3. Morphological class and gender in Arapesh

Arapesh is a language of the Toricelli family, spoken on the north coast of Papua New Guinea, between Dagur and Matapau. Our discussion here draws heavily on Fortune's (1942) grammar and recent work based on it by Aronoff (1992, 1994: 89–114).

Arapesh is a language rich in noun morphological classes and in genders. At least 22 morphological classes can be identified on the basis of noun stem phonology. Table 1 shows the singular/plural phonological alternations characteristic of the ends of noun stems in each of the morphological classes.[3] We follow Aronoff (1992: 31–32) in analyzing the singular forms as bearing no morphological marker, and the plural forms as morphologically complex. Aronoff argues (assuming a process framework) that the plurals result from the action of realization rules which vary according to the morphological class. It is not our concern in this paper to explicate the processes which derive plural forms from stems.

Though the vast majority of nouns can be assigned to one of the morphological classes on the basis of the phonology of their stem alone, a small number of nouns fall outside the system outlined. For example, a few noun stems end in *k*, *b and s*, though no morphological assignments are defined for these endings. According to Fortune, only two nouns in the language end in *b* (*kwagesab* 'croton' and *mib* 'thigh'), and likewise only two end in *k* (*bokok* 'cannibalistic ogre' and *ñibiok* 'sacred flute'). In such cases, the nouns fall into the default morphological class. Nouns belonging to the default class all bear the plural marker *ehas*.

There are also some nouns whose final segment ought to lead them to be assigned to a class by one of the standard assignment rules. Nonetheless, these exceptional cases fail to be assigned in the normal fashion, instead being assigned to the default morphological class and receiving the *ehas* plural marker. An example of this kind is *lim* 'roller for launching canoe', which should be assigned to morphological class 11 and take the plural marker *ipi̥*. Instead, it bears the default class plural marker, and is realised as *limehas*.

A very small number of items do not fit into either the normal or the default assignment systems. Fortune (1942: 14) lists five nouns whose stems end in *n* which ought to have their plurals in *b* (Class 12). Instead, these mark plurality with Class 1 *bys*. So, for example, *awhon*/*awhobys* 'eel'.

[3] Table 1 is adapted from Aronoff (1992: 23); note in particular that the morphological class labels used here differ from those used by Aronoff.

Table 1
Arapesh inflectional noun classes (based on Aronoff, 1992: 23)

Class	Alternation	Singular	Plural	Gloss
1	by̥ /bys	agaby̥	agabys	back
2	bør/ryb	ñibør	ñiryb	belly
3	ag/as	aijag	aijas	leg
4	g/gas	aweg	awegas	seed
5	ku̥ /meb	iloku̥	ilameb	a bird
6	ku̥ /rib	yahaku̥	yaharib	a fruit tree
7	ku̥ /ib	unuku̥	unib	teeth mother
8	ku̥ /guhijer	aniku̥	aniguhijer	rattan species
9	ku̥ /ijer	barahoku̥	barahijer	grand-daughter
10	ku̥ /u	amagoku̥	amagou	fly
11	m/ipi̥	irum	iripi̥	breadfruit
12	n/b	narun	narøb	wave
13	n/m	araman	aramum	man
14	iñ.iš	kobiñ	kobiš	ditch
15	V/Vhas	bode	bodehas	stone axe
16	pu̥ /gwis	barupu̥	barugwis	mountain track
17	pu̥ /s	apapu̥	apas	banana
18	r/guh	jur	juguh	snake
19	t/togu̥	alit	alitogu̥	shelf
20	t/gu̥	nybat	nybagu̥	dog
21	uh/ruh	nauh	naruh	tooth
22	ah/eh	atah	ateh	ear

There are thirteen genders in Arapesh. This claim rests on the assumption that gender is syntactic, and therefore available for agreement, whereas morphological class is not. Arapesh nouns determine, directly or indirectly, the agreement form of adjectives, verbs, and prowords. When the various patterns of agreement are examined (Aronoff, 1992: 22–26), the distinction between the 23 morphological classes (22 regular classes plus the default class) dissolves, leaving behind thirteen agreement classes, or genders. Some of the genders are coextensive with morphological classes (1, 2, 11, 12, 13, 18, 21). In the case of four genders, two morphological classes correspond to a single gender (3 and 4, 14 and 15, 16 and 17, 19 and 20). One gender has no less than six corresponding morphological classes (5, 6, 7, 8, 9, 10). Correspondences between morphological class and gender are shown in Table 2.

Thus, for the vast majority of nouns in Arapesh, there is an implicational relationship between stem phonology and morphological class and between morphological class and gender, as shown below:

Stem Phonology ⇒ Morphological Class ⇒ Gender

As in the case of morphological class assignment, however, some nouns fall outside the normal gender assignment system and end up being assigned to the default gender, which is gender VIII. Notice that there appears to be a non-arbitrary connection

Table 2
Correspondences between morphological classes and genders in Arapesh

Morphological class	Gender
1	I
2	II
3	III
4	
5	
6	
7	IV
8	
9	
10	
11	V
12	VI
13	VII
14	VIII
15	
16	IX
17	
18	X
19	XI
20	
21	XII
22	XIII
DEFAULT	VIII

between the default status of gender VIII and the fact that nouns in the default morphological class take their plurals in *ehas*, as do normally assigned nouns in morphological class 15.

Nouns which have been assigned to the default morphological class are automatically assigned the default gender. Thus nouns with stems ending in rare segments like *b*, *k* and *s* are assigned to the default morphological class and, as a result, they are then assigned default gender. In similar fashion, nouns which are lexically marked to take a normal singular but an exceptional plural (such as *lim* 'roller for launching canoe') are assigned to the default morphological class and receive default gender.

Membership of the default morphological class is not a prerequisite for assignment of default gender. A number of nouns which belong to a normal class morphologically are lexically marked to receive the default gender. For example, the noun *diliat* 'side post which supports eaves of house', plural *diliatogu̥*, is a normal member of morphological class 19, but instead of being assigned gender XI, it receives default gender. Examples of this kind underscore the fact that the direction of implication runs from phonology to morphological class to gender. With the exception of very odd cases like *gun* 'sago pounder of stone', plural *gunabys* which appear to belong to neither the normal nor the default morphological classes, phonological exceptions imply membership of the default morphological class, but membership of that class does not necessarily imply phonological exceptionality. Likewise, membership of the default morphological class implies default gender, but default gender does not necessarily imply membership of the default morphological class.

Gender systems always have a semantic core, and so we find semantic assignment rules, which take precedence over the above mentioned formal assignments. Most of the nouns which designate female persons are assigned to gender IV, while all and only the nouns which designate male persons and the exclusively male roles of warfare and men's initiation ceremonies are assigned to gender VII. Nouns which denote persons without differentiation of sex cannot be assigned with confidence to either of these genders. Such nouns also receive the default gender, regardless of their morphological class.

4. A Network Morphology account of morphological class assignment

In our Network Morphology account of the Arapesh data, we have focused on the morphological class and gender assignments rather than the phonology. Thus we list the stem for each lexical entry, and we (redundantly) state the phonology of the end of the stem explicitly. In a full account of the data, including the phonology, this redundancy would be removed. The present analysis is motivated by the fact that list processing (the operation required to examine the end of a stem), though possible in the DATR representation language, would tend to obscure what is going on at the level of assignment to morphological class and gender. To keep the presentation clear it is also necessary to simplify grossly what happens in the phonology of plurals. Following Aronoff, we hold that singulars are morphologically simple, plurals are morphologically complex, and phonological processes relate the two; our analysis simplifies somewhat by constructing plurals out of a concatenated plural-stem (not necessarily the same as the singular) and a suffix. We regard these as simplifications at the phonological level alone.

Consider the lexical entry for *aijag* 'leg' shown in (7). (Numbers in square brackets on the right hand side of the page refer to line numbers in Appendix A. These are shown only where necessary for purposes of exposition.)

```
(7) Aijag:
      < > == NOUN                                   [72]
      <gloss> == leg                                [73]
      <phon stem> == aijag                          [74]
      <phon end stem> == ag                         [75]
      <phon pl stem> == aij.                        [76]
```

The <phon end stem> path is the convenient place in which we note the phonology of the end of the stem in lieu of a full phonological component. The <phon pl stem> path is the theoretically suspect version of the stem to which the appropriate but equally suspect plural ending is appended, in coarse simulation of a more sophisticated phonological process. If it is assumed that all phonological information is automatically available, then all paths beginning with phon should be read as having no special theoretical significance.

In our analysis, all noun lexical entries inherit from the NOUN node, shown in (8). A full listing of our Network Morphology analysis of Arapesh nouns is provided in Appendix A. Automatically generated output of the theory is shown in Appendix B.

```
(8) Noun:
    <sem personhood> == non_person                              [1]
    <sem sex> == undifferentiated                               [2]
    <syn gender> == GENDER_ASSIGN:< "<sem personhood>" >        [3]
    <mor sgW == "<phon stem>"                                   [4]
    <mor pl> == "<phon pl stem>" "<phon pl end>"                [5]
    <mor class> ==
     MOR_CLASS_ASSIGN: < "<phon end stem>" "<sem sex>" >        [6]
    <phon pl end> == PLURAL: < "<mor class>" >                  [7]
    <phon pl stem> == "<phon stem>".                            [8]
```

The singular consists of a bare stem, according to [4]. The plural is more complex. Fact [5] describes the plural as consisting of the lexically-determined plural stem and the morphological-class-determined ending (subject to all preceding caveats).

By default, the plural stem is identical to the singular stem [8], though this may be overridden in the lexical entry, as in the case of *aijag* (plural stem 'aij') [76]. The plural ending is obtained by globally evaluating the "<mor class>" path [6] and then evaluating the result at the PLURAL node [7].

A value for "<mor class>" must be found by global evaluation. Since no appropriate fact is found at the original query node (Aijag), the one stored at the NOUN node is accessed [6]. This tells us that the morphological class (<mor class>) can be determined by looking at the values of the <phon end stem> and <sem sex> paths at the MOR_CLASS_ASSIGN node. The <phon end stem> path at the Aijag node (7) supplies the required information, namely '*-ag*' [75]. The <sem sex> information is only crucial for the purposes of differentiating two classes, namely 12 and 13. However, since path descriptions increase in specificity from left to right, sex information can be ignored except where it is relevant. The MOR_CLASS_ASSIGN node is shown in (9).

```
(9) MOR_CLASS_ASSIGN:
    < > == default_class                                        [9]
    <bY> == 1                                                   [10]
    <b0r> == 2                                                  [11]
    <ag> == 3                                                   [12]
    <g> == 4                                                    [13]
    <m> == 11                                                   [14]
    <n> == 12                                                   [15]
```

```
<n male> == 13        [16]
<iN> == 14            [17]
<_V> == 15            [18]
<r> == 18             [19]
<uh> == 21            [20]
<ah> == 22.           [21]
```

The value of the <ag> path at the MOR_CLASS_ASSIGN node is 3 or, to put it another way, *aijag* belongs to morphological class 3. Recall that our quest to find the morphological class of this word began at fact [7] in fragment (8). Now that we know that the class is '3', it is necessary to evaluate it at the PLURAL node is shown in (10).

```
(10) PLURAL:
     < > == ehas       [22]
     <1>== s           [23]
     <2> == ryb        [24]
     <3> == as         [25]
     <4> == as         [26]
     <5> == meb        [27]
     <6> == rib        [28]
     <7> == ib         [29]
     <8> == guhijer    [30]
     <9> == ijer       [31]
     <10> == u         [32]
     <11> == ipl       [33]
     <12> == b         [34]
     <13> == m         [35]
     <14> == iS        [36]
     <15> == has       [37]
     <16> == gwis      [38]
     <17> == s         [39]
     <18> == guh       [40]
     <19> == ogU       [41]
     <20> == gU        [42]
     <21> == ruh       [43]
     <22> == eh.       [44]
```

Thus is simply a look-up table which associates plural endings with morphological classes. Thus, when we evaluate the path <3> at this node, the value which is retrieved is '*-as*'. When combined with the plural stem '*aij*' the resulting plural form is '*aijas*'.

It can be quite challenging to follow the sequence of steps in the derivation of a result value from a Network Morphology theory. Figure 1 shows the inference steps involved in deriving the plural form of aijag. The derivation should be read as follows. Lines beginning with a '>' are queries. The value in such lines is shown as '?' to signify that it is, as yet, unknown. The first such query initiates a sequence of derivational steps to obtain the desired value, if possible, from the theory. All subsequent queries arise automatically as sub-goals in the process of deriving the result value for the original query. Indentation is used to illuminate the relationships between queries. No fact derived or query addressed in the course of deriving a result for some query appears to the left of that query on the page. Sub-queries and all facts following immediately from them are indented to the right. When the result of a query is derived, it is indented leftwards so as to line up vertically with the original query. Result facts are written with a preceding '<' for clarity.[4] Numbering in the rightmost column identifies the origin of each fact involved in the derivation. Arabic numerals identify **static** facts, that is, facts stated explicitly in the theory; Roman numerals identify **dynamic** facts, that is, ephemeral facts created during the course of this derivation. Constructs of the form [c](a,b) are used to identify a fact [c], which has been infered from previously introduced facts [a] and [b] (where [c] is a dynamic fact and [a] and [b] may be facts of either kind).

```
> Aijag:<mor plural> = ?                                               ?
  Aijag:< > == NOUN                                                    [72]
  NOUN:<mor plural> == "<phon pl stem>" "<phon pl end>"                [5]
  Aijag:<mor plural> == "<phon pl stem>" "<phon pl end>"               [i] (72, 5)
  > Aijag:<phon pl stem> = ?                                           ? (i)
  < Aijag:<phon pl stem> = aij                                         [ii] (76)
  > Aijag:<phon pl end> = ?                                            ? (i)
    Aijag:<phon pl end> == PLURAL:< "<mor class>" >                    [iii] (72, 7)
    > Aijag:<mor class> = ?                                            ? (iii)
      Aijag:<mor class> ==
        MOR_CLASS_ASSIGN:< "<phon end stem>" "<sem sex>" >             [iv] (72, 6)
      > Aijag:<phon end stem> = ?                                      ? (iv)
      < Aijag:<phon end stem> = ag                                     [v] (75)
      > Aijag:<sem sex> = ?                                            ? (iv)
      < Aijag:<sem sex> = undifferentiated                             [vi] (72, 2)
      > MOR_CLASS_ASSIGN:<ag unfifferentiated> = ?                     ? (iv, v, vi)
      > MOR_CLASS_ASSIGN:<ag undifferentiated> = 3                     [12]
    < Aijag:<mor class> = 3                                            [vii] (iv, 12)
    > PLURAL:<3> = ?                                                   ? (iii, vii)
    < PLURAL:<3< = as                                                  [25]
  < Aijag:<phon·pl end> = as                                           [viii] (iii, vii, 25)
< Aijag:<mor plural> = aij as                                          [ix] (i, ii, viii)
```

Fig. 1. Derivation of the plural form of *Aijag*

This relatively straightforward set of inference steps can be used to derive the plural forms of a considerable proportion of the nouns of Arapesh. However, automatic assignment is not possible in all cases. Consider nouns in morphological classes 19 and 20, for example. Nouns in both of these classes have stems which end

[4] For reasons which need not detain us here, result facts (or 'theorems') are written with a single equals sign rather than the more familiar double equals (Evans and Gazdar, 1989b).

in *-t*. However, class 19 nouns take their plural in *-togu̥* (for example, *alit/alitogu̥* 'shelf'), while class 20 nouns take their plural in *-gu̥* (for example, *nybat/nybagu̥* 'dog'). In cases such as these, it is necessary to state the morphological class explicitly in the lexical entry, as shown in (11) and (12).

(11) Alit:

```
<> == NOUN
<gloss> == shelf
<mor class> == 19
<phon stem> == alit
<phon end stem> == t.
```

(12) Nybat:

```
<> == NOUN
<gloss> == dog
<mor class> == 20
<phon stem> == nybat
<phon end stem> == t
<phon pl stem> == nyba.
```

Now let us consider classes 12 and 13. Nouns in both classes have stems which end in *-n*. However, class 12 nouns take their plural in *-b* (for example, *narun/narøb* 'wave'), while class 13 nouns take their plural in *-m* (for example, *araman/aramum* 'man'). The crucial difference is semantic: class 13 is made up entirely of nouns with male reference, while no such reference is found in class 12. Example lexical entries are shown in (13) and (14). When an attempt is made to retrieve a <mor class> value for these words [6], the MOR_CLASS_ASSIGN node is looked up with the path <n undifferentiated> in the case of *narun/narøb* and <n male> in the case of *araman/aramum*. The latter is exactly matched by fact [16] at the MOR_CLASS_ASSIGN node (9), resulting in assignment to class 13, while the former is included in the partial description at [15], so it is assigned to class 12. It is true that the pattern <n> at [15] constitutes a valid description of both paths, but this does not result in problems, since the semantics of DATR state that in such cases of conflict, the longest matching path takes priority.

(13) Narun:

```
<> == NOUN
<gloss> == wave
<phon stem> == narun
<phon end stem> == n
<phon pl stem> == naro.
```

(14) Araman:

```
<> == NOUN
<gloss> == man
<sem personhood> == person
```

```
<sem sex> == male
<phon stem> == araman
<phon end stem> == n
<phon pl stem> == aramu.
```

We have noted that a variety of exceptions exist to the general pattern of assignments. Consider the example of *bokok/bokokehas* 'cannibalistic ogre', whose stem ending falls outside the normal assignment pattern. The lexical entry for this word is shown in (15).

(15)
```
Bokok:
    <> == NOUN
    <gloss> == cannibalistic_ogre
    <phon stem> == bokok
    <phon end stem> == k.
```

Since the lexical entry does not explicitly assign the word to a morphological class, the same mechanism as that used for *aijag* is used. However, when an attempt is made to find the non-standard *-k* ending at the MOR_CLASS_ASSIGN node (9), an exact match is not found. Instead, the value stored for the maximally underspecified path '< >' is retrieved, namely default_class. Attempting to retrieve a plural ending at the PLURAL node (10) has a similar effect: there is no path of that name (i.e. default_class) at that node, so the value stored in connection with the maximally underspecified node is retrieved, namely '*-ehas*'.

The logic of this accords exactly with the linguistic facts we wish to capture. Use of the maximally underspecified path emphasises the fact that the default class in Arapesh is the 'elsewhere' class, the class to which words should be assigned when all else fails. Note that though we use an explicit label 'default_class', the analysis works without any explicit assignments to a path with this name. The point is not that we have called it by a special name; rather, it is that we have **not** called it by a special name, i.e. 1, 2, ..., 22.

In the case of lexical exceptions, it is necessary to indicate the exceptionality in the lexical entry, as shown in (16).

(16)
```
Lim:
    <> == NOUN
    <gloss> == launching_roller
    <mor class> == MOR_CLASS_ASSIGN:<undefined>
    <phon stem> == lim
    <phon end stem> == m.
```

The logic of this analysis is as follows. Under normal circumstances the end of stem phonology /m/ would be evaluated at the MOR_CLASS_ASSIGN node and would result in straightforward assignment to class 11. However, since *lim/limehas* 'launching roller' violates this pattern, it is necessary to block normal assignment with an overriding fact in the lexical entry, namely MOR_CLASS_ASSIGN:<undefined>. Once again, undefined is no more than an exceptionality marker. The analysis does not work because this atom is recognised; rather, it works because it is **not** recognised. The account would still go through successfully if undefined were replaced with an arbitrary symbol such as something_funny_here. In other words, the account is capturing the fact that anything which fails to fit the normal pattern ends up being assigned to the same class.

In extreme cases such as that of *gun/gunabys* 'sago pounder' (17), all aspects of morphological identity diverge. The singular stem looks like a well-formed class 12 (or 13) word, but the plural looks like a normal class 1 plural. The morphological class could reasonably be said to be undefined. Not surprisingly, *gun/gunabys* takes the default gender. This exceptionality has to be marked in the lexicon by indicating the plural form and stating that the <mor class> is undefined. This overrides the normal morphological class assignment sequence and sets up an unknown symbol (undefined) to defeat the gender assignment system, resulting in assignment to the default gender.

(17) Gun:
```
    <> == NOUN
    <gloss> == sago_pounder
    <mor pl> == gunabys
    <mor class> == undefined
    <phon stem> == gun
    <phon end stem> == n.
```

5. A Network Morphology account of gender assignment

Most of the function of gender assignment is served by three nodes, GENDER_ASSIGN, GENDER_FROM_SEX, and GENDER_FROM_MOR. GENDER_ASSIGN is used to check whether a lexical item denotes a person or not. If it does, gender assignment is based purely on semantics, specifically, sex. If it does not, gender assignment is based purely on formal criteria, namely morphological class. These nodes are shown below.

(18) GENDER_ASSIGN:
```
    <> == _VIII                                                    [45]
    <person> == GENDER_FROM_SEX: <"<sem sex>" >                    [46]
    <non_person> == GENDER_FROM_MOR:< "<mor class>" >.             [47]
```

(19) GENDER_FROM_SEX:
```
    < > == GENDER_ASSIGN                                           [48]
    <female> == _IV                                                [49]
    <male> == _VII.                                                [50]
```

(20)
```
GENDER_FROM_MOR:
    < > == GENDER_ASSIGN        [51]
    <1> == _I                   [52]
    <2> == _II                  [53]
    <3> == _III                 [54]
    <4> == _III                 [55]
    <5> == _IV                  [56]
    <6< == _IV                  [57]
    <7> == _IV                  [58]
    <8> == _IV                  [59]
    <9> == _IV                  [60]
    <10> == _IV                 [61]
    <11> == _V                  [62]
    <12> == _VI                 [63]
    <13> == _VII                [64]
    <16> == _IX                 [65]
    <17> == _IX                 [66]
    <18> == _X                  [67]
    <19> == _XI                 [68]
    <20> == _XI                 [69]
    <21> == _XII                [70]
    <22> == _XIII.              [71]
```

In order to find the gender of a noun, the 'personhood' of the noun's denotatum must be evaluated at the GENDER_ASSIGN node. Personhood is not mentioned in the lexical entry for *aijag* (7). However, all nouns inherit from the NOUN node, where it is stated that the default personhood for a noun is 'non-person' [1]. Thus for *aijag* personhood is 'non-person'. In order to find a value for the path <non_person> at node GENDER_ASSIGN, it is necessary to evaluate <mor class> [47]. We have already learned that the morphological class of *aijag* is 3. Evaluating <3> at the node GENDER_FROM_MOR yields the result III [54]. We have established that *aijag* takes gender III on the basis of reasoning from the phonology of the stem, to the morphological class, to the gender.

Now consider, a lexical entry denoting a person, such as *barahoku/barahoguhijer* 'grand-daughter' (21).

(21)
```
BarahokU:
    <> == NOUN
    <gloss> == grand_daughter
    <sem personhood> == person
    <sem sex> == female
    <mor class> == 9
    <phon stem> == barahokU
    <phon end stem> == kU
    <phon pl stem> == barah.
```

Since the lexical entry tells us that <sem personhood> == person, we are faced with a different option at GENDER_ASSIGN, namely the requirement to evaluate <sem sex> at GENDER_FROM_SEX [46]. The lexical entry tells us that <sem sex> == female, so we can use this information to retrieve the gender IV from the node GENDER_FROM_SEX [49].

What happens if a word does not refer to a person, but it belongs to the default morphological class, as in the case of *ñibiok/ñibiokehas* 'sacred flute' (22)? As with all words which denote non-persons, derivation ends up at the GENDER_FROM_MOR node [47]. This node is a simple table of mappings from morphological classes to genders, very similar to that shown in Table 2. However, there is no explicit mapping from 'default_class' to anything. Instead, there is a default mapping from the maximally underspecified path '< >' back to the GENDER_ASSIGN node where another local default maps the maximally underspecified path to gender VIII, the default gender. In this way, any word which falls outside the normal assignment system ends up being assigned the default gender.

(22)
```
Nibiok:
    <> == NOUN
    <gloss> == sacred_flute
    <phon stem> == _Nibiok
    <phon end stem> == k.
```

In cases where the morphology is regular, but the gender is not, it is necessary to mark the exceptionality lexically. *Diliat/diliatogu̥* 'eave support' is such a case (23). Here, normal gender assignment is blocked by the presence of an overriding fact in the lexical entry <syn gender> == GENDER_ASSIGN:<undefined>. This captures the intuition that the normal inputs to gender assignment cannot be used. Looking up undefined at GENDER_ASSIGN fails to produce an exact match, so the default value is retrieved, namely VIII, the default gender [45].

(23)
```
Diliat:
    <> == NOUN
    <gloss> == eave_support
    <syn gender> == GENDER_ASSIGN:<undefined>
    <mor class> == 19
    <phon stem> == diliat
    <phon end stem> == t.
```

No special problems for gender assignment are raised by words with extremely irregular morphology such as *gun/gunabys* 'sago pounder', since their failure to belong to any known morphological class ensures that they are assigned to the default gender [51], [45].

6. Coherence of the notion 'default'

In this paper we have taken Aranoff's analysis of Arapesh nouns and demonstrated its coherence and validity by rendering it in a formally rigorous fashion so as to produce the predicted forms. We have done so using the DATR representation language and a minimal set of Network Morphology assumptions whose existence we have already motivated elsewhere (Fraser and Corbett, 1995). Central to our account has been the idea of a default value, that is, a general purpose value which becomes available when no more specific value is available.

The principle of distinguishing between general and exceptional cases has a long and honourable tradition in linguistics, going back at least as far as Pāṇini. In modern linguistics, Kiparsky's (1973) formulation of the Elsewhere Condition did much to establish the idea (compare the closely related Proper Inclusion Precedence Principle of Koutsoudas et al., 1974). Over a period of almost three decades, Stephen Anderson has been developing morphological analyses which depend heavily on disjunctive ordering (Anderson, 1969, 1986, 1992). In syntax, Hudson's Word Grammar (1984, 1990) structures the entire lexicon as a default inheritance hierarchy. Defaults also played an important role in Generalised Phrase Structure Grammar (Gazdar et al., 1985), where the problems associated with interdependent defaults – when a default for a particular feature depends on the value of some other feature – were addressed explicitly and in considerable detail.

In two important studies, Zwicky (1986, 1989) has done much to clarify the notion 'default' in linguistics and to show how it relates to certain other ideas. In particular, he has made explicit the contrast between the 'General as Default' view and the 'General as Basic' view, the latter of which has dominated all branches of generative grammar. In the General as Basic view 'one case is taken to be basic, deep, or underlying; its rule applies first, and another rule alters the basic forms for another case' (Zwicky, 1986: 305). In the General as Default view, there is no possibility of transforming one form into another. Rather, there is a fixed but disjunctively ordered set of alternatives; there is a structured list of exceptions leading to an ultimate default.

In our work on Network Morphology, we have given careful attention to the nature and utility of default representations. However, it was only while working on the implementation of Arapesh noun morphology reported here that we came to realise that the term 'default' – as used by us and by many other linguists – was less coherent than we had thought.

This can be illustrated from the preceding discussion in this paper, in which we have used the term 'default' in two ways which, though conceptually related, are nonetheless formally distinct. In the first case we described

```
"<singular>" s
```

as the default schema for plural formation in English. Since this information is lodged at a node high in the inheritance structure (the COUNT_NOUN node), it will be inherited as a default value by all count nouns which do not specify a plural form at a more specific node. A similar kind of default was provided by the definition of non-person as the default value for the <sem personhood> path at the NOUN node in our analysis of Arapesh. Since the vast majority of nouns denote non-persons, the default personhood can be made available for the relevant nouns simply by leaving the <sem personhood> path unspecified in the lexical entries for those nouns; default inheritance ensures that the default value finds its way into each underspecified lexical entry. We shall call instances of this type **normal case defaults**.

In the second use of the term, a default is something which applies when the normal system breaks down. It is a safety net. Thus, the default morphological class is invoked when some idiosyncratic feature of a lexical entry gets in the way of normal class assignment. Nouns are assigned to the default morphological class or the default gender when something about them interferes with the normal assignment process. We shall call instances of this type **exceptional case defaults**.

There is a common conceptual core running through both usages of the term: the default is the last thing you get to. This is where the similarity ends. A normal case default is retrieved after failing to find any more specific value; an exceptional case default is retrieved after finding too much information – information which blocks normal retrieval and causes a backstop value to be accessed instead. One form of default is concerned with typicality, the other with exceptionality. It is therefore particularly important that conceptual and terminological confusion be avoided by adequate definition of which sense of the term is intended on any given occasion.[5]

One of the payoffs of working in a formally explicit framework such as Network Morphology is that it lays bare the differences between these otherwise confusable notions. In instances where normal case defaults apply, lexical entries are characterised by their brevity. Because a word is fairly typical, many of its parts can be left underspecified, to be filled in by default inheritance. On the other hand, there is an inverse correlation between radically underspecified lexical entries and exceptional case defaults. An exceptional case default is very unlikely to apply unless the lexical entry includes some idiosyncratic information. In our analysis, lexical exceptions such as lim, *gun* and *diliat* include an explicit marker (we use undefined, though any arbitrary extra-theoretic string would suffice) to block inheritance of the normal case default and force invocation of the exceptional case default.

We have confirmed Aronoff's analysis of the Arapesh data, and by expressing it in the DATR formalism (Appendix A) we have been able to demonstrate that the analysis does indeed account for the data, as shown by the output given in Appendix B. Our formal approach has also led to the conclusion that the term 'default' is currently used with two rather different meanings, which linguists would do well to distinguish.

---

[5] For another study which includes consideration of what we term an exceptional case default see Marcus et al. (1993), where *-s* is discussed as the default plural of German. They quote Van Dam (1940), who called it the *Notpluralendung* 'emergency plural ending'. See also Clahsen et al. (1992).

Appendix A

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                                                      %
% File:      apesh12.dtr                               %
% Purpose:   The noun morphology of Arapesh            %
% Authors:   Norman Fraser & Greville Corbett, April 30, 1997 %
% Address:   LIS, University of Surrey, Guildford, Surrey GU2 5XH %
% Version:   1.12                                      %
%                                                      %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% The noun morphology of Arapesh (a language of the Torricelli family, spoken near the
% north coast of Papua New Guinea). The data and part of the analysis are taken from:
%
%   Aronoff, Mark (1991) Noun classes in Arapesh. Yearbook of Morphology, 21–32.
%
%   which, in turn, draws heavily on:
%
%   Fortune, Reo (1942) Arapesh. Publications of the American Ethnological Society XIX.
%     New York: J.J. Augustin.
%
% Most DATR interpreters support only a restricted set of characters, so the following cor-
% respondences have been used:
%
%   Y = y̥      U = u̥
%   0 = ø̊      I = i̥
%   N = ñ      s = ʃ
%
% In addition, the DATR interpreter accords symbols beginning with an upper case letter a
% special status. To overcome this, the '_' (underscore) character is used. It has no theoret-
% ical significance and may safely be ignored.
%
%   All material to the right of a '%' symbol has the status of a comment – it does not
% belong to the substance of the theory.
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

NOUN:
   <sem personhood> == non_person                                         % [1]
   <sem sex> == undifferentiated                                          % [2]
   <syn gender> == GENDER_ASSIGN:< "<sem personhood>" >                   % [3]
   <mor sg> == "<phon stem>"                                              % [4]
   <mor pl> == "<phon pl stem>" "<phon pl end>"                           % [5]
   <mor class> == MOR_CLASS_ASSIGN:< "<phon end stem>" "<sem sex>" >      % [6]
   <phon pl end> == PLURAL:< "<mor class>" >                              % [7]
   <phon pl stem> == "<phon stem>".                                       % [8]
```

```
MOR_CLASS_ASSIGN:
   < > == default_class                                        % [9]
   <bY> == 1                                                   % [10]
   <b0r> == 2                                                  % [11]
   <ag> == 3                                                   % [12]
   <g> == 4                                                    % [13]
   <m> == 11                                                   % [14]
   <n> == 12                                                   % [15]
   <n male> == 13                                              % [16]
   <iN> == 14                                                  % [17]
   <_V> == 15                                                  % [18]
   <r> == 18                                                   % [19]
   <uh> == 21                                                  % [20]
   <ah> == 22.                                                 % [21]

   PLURAL:
   < > == ehas                                                 % [22]
   <1> == bys                                                  % [23]
   <2> == ryb                                                  % [24]
   <3> == as                                                   % [25]
   <4> == as                                                   % [26]
   <5> == meb                                                  % [27]
   <6> == rib                                                  % [28]
   <7> == ib                                                   % [29]
   <8> == guhijer                                              % [30]
   <9> == ijer                                                 % [31]
   <10> == u                                                   % [32]
   <11> == ipI                                                 % [33]
   <12> == b                                                   % [34]
   <13> == m                                                   % [35]
   <14> == iS                                                  % [36]
   <15> == has                                                 % [37]
   <16> == gwis                                                % [38]
   <17> == s                                                   % [39]
   <18> == guh                                                 % [40]
   <19> == ogU                                                 % [41]
   <20> == gU                                                  % [42]
   <21> == ruh                                                 % [43]
   <22> == eh.                                                 % [44]

 GENDER_ASSIGN:
   < > == _VIII                                                % [45]
   <person> == GENDER_FROM_SEX:< "<sem sex>" >                 % [46]
   <non_person> == GENDER_FROM_MOR:< "<mor class>" >.          % [47]

 GENDER_FROM_SEX:
   < > == GENDER_ASSIGN                                        % [48]
   <female> == _IV                                             % [49]
   <male> == _VII.                                             % [50]
```

```
GENDER_FROM_MOR:
   < > == GENDER_ASSIGN                                          % [51]
   <1> == _I                                                     % [52]
   <2> == _II                                                    % [53]
   <3> == _III                                                   % [54]
   <4> == _III                                                   % [55]
   <5> == _IV                                                    % [56]
   <6> == _IV                                                    % [57]
   <7> == _IV                                                    % [58]
   <8> == _IV                                                    % [59]
   <9> == _IV                                                    % [60]
   <10> == _IV                                                   % [61]
   <11> == _V                                                    % [62]
   <12> == _VI                                                   % [63]
   <13> == _VII                                                  % [64]
   <16> == _IX                                                   % [65]
   <17> == _IX                                                   % [66]
   <18> == _X                                                    % [67]
   <19> == _XI                                                   % [68]
   <20> == _XI                                                   % [69]
   <21> == _XII                                                  % [70]
   <22> == _XIII.                                                % [71]

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                      LEXICAL ENTRIES                       %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

AgabY:                                   % Regular Morphological Class 1
   < > == NOUN
   <gloss> == back
   <phon stem> == agabY
   <phon end stem> == bY
   <phon pl stem> == aga.

Nib0r:                                   % Regular Morphological Class 2
   < > == NOUN
   <gloss> == belly
   <phon stem> == _Nib0r
   <phon end stem> == b0r
   <phon pl stem> == _Ni.

Aijag:                                   % Regular Morphological Class 3
   < > == NOUN                                                   % [72]
   <gloss> == leg                                                % [73]
   <phon stem> == aijag                                          % [74]
   <phon end stem> == ag                                         % [75]
   <phon pl stem> == aij.                                        % [76]
```

```
Aweg:                                    % Regular Morphological Class 4
   < > == NOUN
   <gloss> == seed
   <phon stem> == aweg
   <phon end stem> == g.

IlokU:                                   % Regular Morphological Class 5
   < > == NOUN
   <gloss> == bird
   <mor class> == 5
   <phon stem> == ilokU
   <phon end stem> == kU
   <phon pl stem> == ila.

YahakU:                                  % Regular Morphological Class 6
   < > == NOUN
   <gloss> == fruit_tree
   <mor class> == 6
   <phon stem> == yahakU
   <phon end stem> == kU
   <phon pl stem> == yaha.

UnukU:                                   % Regular Morphological Class 7
   < > == NOUN
   <gloss> == teeth_mother
   <mor class> == 7
   <phon stem> == unukU
   <phon end stem> == kU
   <phon pl stem> == un.

AnikU:                                   % Regular Morphological Class 8
   < > == NOUN
   <gloss> == rattan_species
   <mor class> == 8
   <phon stem> == anikU
   <phon end stem> == kU
   <phon pl stem> == ani.

BarahokU:                                % Regular Morphological Class 9
   < > == NOUN
   <gloss> == grand_daughter             % female person
   <sem personhood> == person
   <sem sex> == female
   <mor class> == 9
   <phon stem> == barahokU
   <phon end stem> == kU
   <phon pl stem> == barah.
```

```
AmagokU:                              % Regular Morphological Class 10
    < > == NOUN
    <gloss> == fly
    <mor class> == 10
    <phon stem> == amagokU
    <phon end stem> == kU
    <phon pl stem> == amago.

Irum:                                 % Regular Morphological Class 11
    < > == NOUN
    <gloss> == breadfruit
    <phon stem> == irum
    <phon end stem> == m
    <phon pl stem> == ir.

Narun:                                % Regular Morphological Class 12
    < > == NOUN
    <gloss> == wave
    <phon stem> == narun
    <phon end stem> == n
    <phon pl stem> == nar0.

Araman:                               % Regular Morphological Class 13
    < > == nOUN
    <gloss> == man
    <sem personhood> == person
    <sem sex> == male
    <phon stem> == araman
    <phon end stem> == n
    <phon pl stem> == aramu.

KobiN:                                % Regular Morphological Class 14
    < > == NOUN
    <gloss> == ditch
    <phon stem> == kobiN
    <phon end stem> == iN
    <phon pl stem> == kob.

Bode:                                 % Regular Morphological Class 15
    < > == NOUN
    <gloss> == stone_axe
    <phon stem> == bode
    <phon end stem> == _V.

BarupU:                               % Regular Morphological Class 16
    < > == NOUN
    <gloss> == mountain_track
    <mor class> == 16
    <phon stem> == barupU
```

```
    <phon end stem> == pU
    <phon pl stem> == baru.

ApapU:                                  % Regular Morphological Class 17
    < > == NOUN
    <gloss> == banana
    <mor class> == 17
    <phon stem> == apapU
    <phon end stem> == pU
    <phon pl stem> == apa.

Jur:                                    % Regular Morphological Class 18
    < > == NOUN
    <gloss> == snake
    <phon stem> == jur
    <phon end stem> == r
    <phon pl stem> == ju.

Alit:                                   % Regular Morphological Class 19
    < > == NOUN
    <gloss> == shelf
    <mor class> == 19
    <phon stem> == alit
    <phon end stem> == t.

Nybat:                                  % Regular Morphological Class 20
    < > == NOUN
    <gloss> == dog
    <mor class> == 20
    <phon stem> == nybat
    <phon end stem> == t
    <phon pl stem> == nyba.

Nauh:                                   % Regular Morphological Class 21
    < > == NOUN
    <gloss> == tooth
    <phon stem> == nauh
    <phon end stem> == uh
    <phon pl stem> == na.

Atah:                                   % Regular Morphological Class 22
    < > == NOUN
    <gloss> == ear
    <phon stem> == atah
    <phon end stem> == ah
    <phon pl stem> == at.

Kwagesab:                               % Default Morphological Class
    < > == NOUN
```

```
    <gloss> == croton
    <phon stem> == kwagesab
    <phon end stem> == b.                  % exceptional phonology

Mib:                                       % Default Morphological Class
    < > == NOUN
    <gloss> == thigh
    <phon stem> == mib
    <phon end stem> == b.                  % exceptional phonology

Bokok:                                     % Default Morphological Class
    < > == NOUN
    <gloss> == cannibalistic_ogre
    <phon stem> == bokok
    <phon end stem> == k.                  % exceptional phonology

Nibiok:                                    % Default Morphological Class
    < > == NOUN
    <gloss> == sacred_flute
    <phon stem> == _Nibiok
    <phon end stem> == k.                  % exceptional phonology

Pas:                                       % Default Morphological Class
    < > == NOUN
    <gloss> == taro_pounder
    <phon stem> == pas
    <phon end stem> == s.                  % exceptional phonology

Lim:                                       % Default Morphological Class
    < > == NOUN
    <gloss> == launching_roller            % Lexical exception – ought to
    <mor class> == MOR_CLASS_ASSIGN:<undefined>
    <phon stem> == lim                     % belong to class 11
    <phon end stem> == m.                  % but doesn't

Sam:                                       % Default Morphological Class
    < > == NOUN
    <gloss> == croquettes                  % Another lexical exception.
    <mor class> == MOR_CLASS_ASSIGN:<undefined>
    <phon stem> == sam                     % Should be in class 11,
    <phon end stem> == m.                  % but isn't.

Diliat:                                    % Regular Morphological Class
    < > == NOUN                            % Default gender
    <gloss> == eave_support
    <syn gender> == GENDER_ASSIGN:<undefined>  % Use normal gender
    <mor class> == 19                      % assignment method with undefined
    <phon stem> == diliat                  % input.
    <phon end stem> == t.
```

```
Gun:                                    % Completely exceptional
   < > == NOUN
   <gloss> == sago_pounder              % sg phon => mor class 12
   <mor pl> == gunabys                  % pl phon => mor class 1
   <mor class> == undefined             % Overrides normal morphological
   <phon stem> == gun                   % class assignment.
   <phon end stem> == n.
```

Appendix B

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                                          %
%        Output of apesh12.dtr 30/4/97     %
%                                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

AgabY:<gloss> = back.
AgabY:<syn gender> = _I.
AgabY:<mor class> = 1.
AgabY:<mor sg> = agabY.
AgabY:<mor pl> = aga bys.

Nib0r:<gloss> = belly.
Nib0r:<syn gender> = _II.
Nib0r:<mor class> = 2.
Nib0r:<mor sg> = _Nib0r.
Nib0r:<mor pl> = _Ni ryb.

Aijag:<gloss> = leg.
Aijag:<syn gender> = _III.
Aijag:<mor class> = 3.
Aijag:<mor sg> = aijag.
Aijag:<mor pl> = aij as.

Aweg:<gloss> = seed.
Aweg:<syn gender> = _III.
Aweg:<mor class> = 4.
Aweg:<mor sg> = aweg.
Aweg:<mor pl> = aweg as.

IlokU:<gloss> = bird.
IlokU:<syn gender> = _IV.
IlokU:<mor class> = 5.
IlokU:<mor sg> = ilokU.
IlokU:<mor pl> = ila meb.

YahakU:<gloss> = fruit_tree.
YahakU:<syn gender> = _IV.
```

YahakU:<mor class> = 6.
YahakU:<mor sg> = yahakU.
YahakU:<mor pl> = yaha rib.

UnukU:<gloss> = teeth_mother.
UnukU:<syn gender> = _IV.
UnukU:<mor class> = 7.
UnukU:<mor sg> = unukU.
UnukU:<mor pl> = un ib.

AnikU:<gloss> = rattan_species.
AnikU:<syn gender> = _IV.
AnikU:<mor class> = 8.
AnikU:<mor sg> = anikU.
AnikU:<mor pl> = ani guhijer.

BarahokU:<gloss> = grand_daughter.
BarahokU:<syn gender> = _IV.
BarahokU:<mor class> = 9.
BarahokU:<mor sg> = barahokU.
BarahokU:<mor pl> = barah ijer.

AmagokU:<glosss> = fly.
AmagokU:<syn gender> = _IV.
AmagokU:<mor class> = 10.
AmagokU:<mor sg> = amagokU.
AmagokU:<mor pl> = amago u.

Irum:<gloss> = breadfruit.
Irum:<syn gender> = _V.
Irum:<mor class> = 11.
Irum:<mor sg> = irum.
Irum:<mor pl> = ir ipI.

Narun:<gloss> = wave.
Narun:<syn gender> = _VI.
Narun:<mor class> = 12.
Narun:<mor sg> = narun.
Narun:<mor pl> = nar0 b.

Araman:<gloss> = man.
Araman:<syn gender> = _VII.
Araman:<mor class> = 13.
Araman:<mor sg> = araman.
Araman:<mor pl> = aramu m.

KobiN:<gloss> = ditch.
KobiN:<syn gender> = _VIII.
KobiN:<mor class> = 14.

Kobin:<mor sg> = kobiN.
KobiN:<mor pl> = kob iS.

Bode:<gloss> = stone_axe.
Bode:<syn gender> = _VIII.
Bode:<mor class> = 15.
Bode:<mor sg> = bode.
Bode:<mor pl> = bode has.

BarupU:<gloss> = mountain_track.
BarupU:<syn gender> = _IX.
BarupU:<mor class> = 16.
BarupU:<mor sg> = barupU.
BarupU:<mor pl> = baru gwis.

ApapU:<gloss> = banana.
ApapU:<syn gender> = _IX.
ApapU:<mor class> = 17.
ApapU:<mor sg> = apapU.
ApapU:<mor pl> = apa S.

Jur:<gloss> = snake.
Jur:<syn gender> = _X.
Jur:<mor class> = 18.
Jur:<mor sg> = jur.
Jur:<mor pl> = ju guh.

Alit:<gloss> = shelf.
Alit:<syn gender> = _XI.
Alit:<mor class> = 19.
Alit:<mor sg> = alit.
Alit:<mor pl> = alit ogU.

Nybat:<gloss> = dog.
Nybat:<syn gender> = _XI.
Nybat:<mor class> = 20.
Nybat:<mor sg> = nybat.
Nybat:<mor pl> = nyba gU.

Nauh:<gloss> = tooth.
Nauh:<syn gender> = _XII.
Nauh:<mor class> = 21.
Nauh:<mor sg> = nauh.
Nauh:<ùpŕˆm> = na ruh.

Atah:<gloss> = ear.
Atah:<syn gender> = _XIII.
Atah:<mor class> = 22.
Atah:<mor sg> = atah.

Atah:<mor pl> = at eh.

Kwagesab:<gloss> = croton.
Kwagesab:<syn gender> = _VIII.
Kwagesab:<mor class> = default_class.
Kwagesab:<mor sg> = kwagesab.
Kwagesab:<mor pl> = kwagesab ehas.

Mib:<gloss> = thigh.
Mib:<syn gender> = _VIII.
Mib:<mor class> = default_class.
Mib:<mor sg> = mib.
Mib:<mor pl> = mib ehas.

Bokok:<gloss> = cannibalistic_ogre.
Bokok:<syn gender> = _VIII.
Bokok:<mor class> = default_class.
Bokok:<mor sg> = bokok.
Bokok:<mor pl> = bokok ehas.

Nibiok:<gloss> = sacred_flute.
Nibiok:<syn gender> = _VIII.
Nibiok:<mor class> = default_class.
Nibiok:<mor sg> = _Nibiok.
Nibiok:<mor pl> = _Nibiok ehas.

Pas:<gloss> = taro_pounder.
Pas:<syn gender> = _VIII.
Pas:<mor class> = default_class.
Pas:<mar sg> = pas.
Pas:<mor pl> = pas ehas.

Lim:<gloss> = launching_roller.
Lim:<syn gender> = _VIII.
Lim:<mor class> = default_class.
Lim:<mor sg> = lim.
Lim:<mor pl> = lim ehas.

Sam:<gloss> = croquettes.
Sam:<syn gender> = _VIII.
Sam:<mor class> = default_class.
Sam:<mor sg> = sam.
Sam:<mor pl> = sam ehas.

Diliat:<gloss> = eave_support.
Diliat:<syn gender> = _VIII.
Diliat:<mor class> = 19.
Diliat:<mor sg> = diliat.
Diliat:<mor pl> = diliat ogU.

Gun:<gloss> = sago_pounder.
Gun:<syn gender> = _VIII.
Gun:<mor class> = undefined.
Gun:<mor sg> = gun.
Gun:<mor pl> = gunabys.

References

Anderson, S.R., 1969. West Scandanavian vowel systems and the ordering of phonological rules. Doctoral dissertation, MIT.

Anderson, S.R., 1986. Disjunctive ordering in inflectional morphology. Natural Language and Linguistic Theory 4, 1–32.

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

Aronoff, M., 1992. Noun classes in Arapesh. In: G. Booij, J. van Marle (eds.), Yearbook of Morphology 1991, 21–32. Dordrecht: Kluwer.

Aronoff, M., 1994. Morphology by itself: Stems and inflectional classes. Linguistic Inquiry Monograph 22. Cambridge, MA: MIT Press.

Brown, D., G.G. Corbett, N.M. Fraser, A. Hippisley and A. Timberlake, 1996. Russian noun stress and Network Morphology. Linguistics 34, 52–107.

Brown, D. and A. Hippisley, 1994. Conflict in Russian genitive plural assignment: A solution implemented in DATR. Journal of Slavic Linguistics 2, 48–76.

Clahsen, H., M. Rothweiler, A. Woest and G.F. Marcus, 1992. Regular and irregular inflection in the acquisition of German noun plurals. Cognition 45, 225-255.

Corbett, G.G., 1991. Gender. Cambridge: Cambridge University Press.

Corbett, G.G. and N.M. Fraser, 1993. Network morphology: A DATR account of Russian inflectional morphology. Journal of Linguistics 29, 113–142.

Evans, R. and G. Gazdar, 1989a. The semantics of DATR. In: A.G. Cohn (ed.), Proceedings of the Seventh Conference of the Society for the Study of Artificial Intelligence and Simulation of Behaviour, 79–87. London: Pitman/Morgan Kaufmann.

Evans, R. and G. Gazdar, 1989b. Inference in DATR. In: Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics, 66–71. Manchester: Association for Computational Linguistics.

Evans, R. and G. Dazdar, 1996. DATR: A language for lexical knowledge representation. Computational Linguistics 22; 167–216.

Fortune, R.F., 1942. Arapesh. Publications of the American Ethnological Society XIX. New York: Augustin.

Fraser, N.M., 1985. A Word Grammar parser. Master's dissertation, University College London.

Fraser, N.M. and G.G. Corbett, 1995. Gender, animacy, and declensional class assignment: A unified account for Russian. In: G. Booij, J. van Marle (eds.), Yearbook of Morphology 1994, 123–150. Dordrecht: Kluwer.

Gazdar, G., forthcoming. Ceteris paribus. To appear in: J.A.W. Kamp, C. Rohrer (eds.), Aspects of computational linuistics. Berlin: Springer.

Gazdar, G., E. Klein, G. Pullum and I. Sag, 1985. Generalized Phrase Structure Grammar. Oxford: Blackwell.

Gibbon, D., 1992. ILEX: A linguistic approach to computational lexica. In: U. Klenk (ed.), Computatio Linguae: Aufsätze zur algorithmischen und quantitativen Analyse der Sprache (= Zeitschrift für Dialektologie und Linguistik, Beiheft 73), 32–53. Stuttgart: Franz Steiner.

Hudson, R.A., 1984. Word Grammar. Oxford: Blackwell.

Hudson, R.A., 1990. English Word Grammar. Oxford: Blackwell.

Kiparsky, P., 1973. 'Elsewhere' in phonology. In: S.R. Anderson, P. Kiparsky (eds.), A Festschrift for Morris Halle, 93–106. New York: Holt, Rubegart, and Winston.

Koutsoudas, A., G. Sanders and C. Noll, 1974. On the application of phonological rules. Language 50, 1–28.

Marcus, G.F., U. Brinkmann, H. Clahsen, R. Wiese, A. Woest and S. Pinker, 1993. German inflection: The exception that proves the rule. Theorie des Lexikons: Arbeiten des Sonderforschungsbereichs 282, 40. Düsseldorf: Heinrich-Heine-Universität, Seminar für Allgemeine Sprachwissenschaft.

Pullum, G.K., 1989. The incident of the node vortex problem. Natural Language and Linguistic Theory 3, 473–479.

Van Dam, J., 1940. Handbuch der deutschen Sprache: Zweiter Band: Wortlehre. Groningen: Wolters. [Quoted from Marcus, Brinkmann, Clahsen, Wiese, Woest and Pinker, 1993.]

Zwicky, A., 1986. The general case: Basic form versus default form. In: V. Nikiforidou, M. VanClay, M. Niepokuj, D. Feder (eds.), Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society, 305–314. Berkeley, CA: Berkeley Linguistic Society.

Zwicky, A., 1989. What's become of derivations? Defaults and invocations. In: K. Hall, M. Meacham, R. Shapiro (eds.), Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society, 303–320. Berkeley, CA: Berkeley Linguistic Zociety.