Where word forms collide: a typology of syncretism
Full Research Report

1 Background
Syncretism is a surprising, widespread and yet poorly understood phenomenon in natural language. Given a regular distinction such as present versus past, as in English help/helped, work/worked, laugh/laughed, we might not expect to find instances like bid, which can be present or past (we now bid five pounds, though yesterday we bid ten pounds). The form bid, is said to be an instance of syncretism, a single form fulfilling two different functions. Thus syncretism is found even in English, whose inflectional morphology (system of different word-forms) is simple in comparison with many languages. Though syncretisms are found frequently, and linguists often refer to the notion, little was previously known about their typological and geographical distribution, and about the range of patterns attested in the world’s languages, or any universal constraints governing possible patterns. Our project has addressed all of these issues in depth: we have contributed publications on typological and geographical distribution of syncretism; our detailed database study has allowed us to map out the range of patterns; and we have made proposals, founded in sound empirical work, about the different constraints underlying syncretism, thereby providing a firm foundation for the general theory of syncretism which we will publish in the Cambridge University Press Blue Series.

2 Objectives
There were three main objectives:

2.1 To create a ‘morphological diversity database’ from a typologically broad set of languages. We created a comprehensive database of syncretism in thirty languages (see §3, §4.1). In addition to that, we created a second database covering syncretism of subject person in verbs in 111 languages. Both databases are available online at <www.smg.surrey.ac.uk>.

2.2 To define criteria for identifying instances of syncretism. We attributed the surface effect of syncretism to three formally distinct causes, and developed criteria for identifying and analysing each of these (see §4.2.2).

2.3 To devise a general theory of syncretism On the foundation of an unprecedented wealth of material, we have been able to describe the typological range of syncretism on the basis of an articulated formal theory and to critically assess previous models (see the discussion throughout §4).

3 Methods
The two relational databases constitute the empirical basis for the theoretical and typological conclusions outlined in §4. The principles of their construction are outlined below.

3.1 Sampling
The purpose of the databases is twofold: (i) to record syncretism in a genetically and geographically diverse sample of languages; (ii) to examine the logical space of syncretism by recording the greatest typological variety. We had to balance these goals, as they are potentially in conflict: genetic diversity does not necessarily correspond to typological diversity. A further complication is created by the very nature of syncretism, which by definition is limited to languages that have inflectional morphology, thus some families, such as Sinitic, were eliminated from consideration at the very outset. Since one of the stated goals was to extend our empirical foundation beyond the Indo-European examples typically cited in discussions of syncretism, these were consciously avoided in the initial sampling. However, given the extraordinary range of syncretic patterns found in Indo-European languages, a relatively high number (four) of these were finally included in order to fill out the logical typological space. The full list of languages in the database is given in figure 1. The sources were published grammars, and in some cases consultation with experts.

<table>
<thead>
<tr>
<th>Family</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afroasiatic</td>
<td>Classical Arabic (West Asia)</td>
</tr>
</tbody>
</table>
For the database on person syncretism, the core sample was taken from that established for the World Atlas of Language Structures project (see §6.4). This was expanded by an additional 46 languages, bringing the total to 111 languages, representing 46 different families; the full list is available at <http://www.smg.surrey.ac.uk>.

3.2 Analysis

Syncretism is identified on the basis of the comparison of the actual repertory of inflected forms with an idealised underlying morphosyntactic paradigm. The morphosyntactic paradigm has been generated as follows: (i) for any language, establish the morphosyntactic features (e.g. number) and their values (e.g. singular, plural) which are correlated with some distinct inflectional behaviour; (ii) for each distinct word class within a language, establish which morphosyntactic features are in operation; and (iii) project all the logically possible feature value combinations to produce the underlying morphosyntactic paradigm. For example, if nouns in a language are determined to bear the features ‘case’ (with the values ‘nominative’, ‘accusative’ and ‘genitive’) and ‘number’ (with values singular, dual and plural), these multiply out into a paradigmatic grid with nine cells (fig. 2). Within that language, any form which unites the values of two or more cells is considered syncretic.

3.3 Structure of the databases

<table>
<thead>
<tr>
<th></th>
<th>NOM SG</th>
<th>NOM DU</th>
<th>NOM PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC SG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEN SG</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Languages in the main database

<table>
<thead>
<tr>
<th>Language</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somali</td>
<td>Africa</td>
</tr>
<tr>
<td>Bao’an</td>
<td>Central Asia</td>
</tr>
<tr>
<td>Austronesian</td>
<td>Pacific</td>
</tr>
<tr>
<td>Carib</td>
<td>South America</td>
</tr>
<tr>
<td>Chibchan</td>
<td>South America</td>
</tr>
<tr>
<td>Chukotko-Kamchatkan</td>
<td>Siberia</td>
</tr>
<tr>
<td>Dravidian</td>
<td>South Asia</td>
</tr>
<tr>
<td>Eskimo-Aleut</td>
<td>North America</td>
</tr>
<tr>
<td>Indo-European</td>
<td>Caucasus</td>
</tr>
<tr>
<td>Isolate</td>
<td>South Asia</td>
</tr>
<tr>
<td>Kartvelian</td>
<td>Caucasus</td>
</tr>
<tr>
<td>Kiowa-Tanoan</td>
<td>North America</td>
</tr>
<tr>
<td>Nakh-Daghestanian</td>
<td>Caucasus</td>
</tr>
<tr>
<td>Niger-Kordofanian</td>
<td>Africa</td>
</tr>
<tr>
<td>Nilo-Saharan</td>
<td>Africa</td>
</tr>
<tr>
<td>Non-Pama-Nyungan</td>
<td>Australia</td>
</tr>
<tr>
<td>Australian</td>
<td>Mayali</td>
</tr>
<tr>
<td>Pama-Nyungan</td>
<td>Australia</td>
</tr>
<tr>
<td>Sepik-Ramu</td>
<td>Pacific</td>
</tr>
<tr>
<td>Tacanan</td>
<td>South America</td>
</tr>
<tr>
<td>Tibeto-Burman</td>
<td>South Asia</td>
</tr>
<tr>
<td>Trans-New Guinea</td>
<td>Pacific</td>
</tr>
<tr>
<td>Uralic</td>
<td>Europe</td>
</tr>
<tr>
<td>Erzja Mordvin</td>
<td>Europe</td>
</tr>
</tbody>
</table>
The main database contains every instance of syncretism from a sample of 30 languages. The other database focuses on person syncretism in 111 languages. All of the data entered in these two databases is available over the web with an easy to use interface (see §5 and §6). In this section we describe the underlying databases from which the web versions were created. A user does not need to know this information to use the web version, which is somewhat simpler, and is accompanied by clearly written readme files.

3.3.1 Main database

The main database is written in Microsoft Access, and contains eighteen tables. Of these, ten tables contain the values for well-defined feature sets: number, case, gender, definiteness, person, tense, mood, voice, aspect and negation. These tables are to be seen on the right side of figure 3.

![Database Relationships](image)

Figure 3: Database Relationships

A feature value from one of these tables can be combined with values of any of the other features to form a morphosyntactic combination, e.g. ‘2sg past’ is a combination of values of the features person, number and tense. The Combination table is ringed in figure 3. The relationship between any given feature table and the Combination table is one-to-many.

The table of syncretisms is in the middle of the top set of tables. The Combination table has two relationships with the Syncretism table. Both of these are one-to-many. A syncretism is treated as a binary pair of morphosyntactic combinations. The Syncretism table contains the indices for each morphosyntactic combination. These indices were automatically compared during data entry to check that they are not identical (as a form which shares identical morphosyntactic features is not a syncretism). Other fields in the LanguageSyncretismDomain table (see below) allowed us to associate a binary pair with other binary pairs, if they were part of a greater syncretism. This was done by defining a syncretism as ‘independent’ if it could occur on its own as an independent pattern in the language, and ‘dependent’ if the binary pair only occurs in the presence of another syncretism. A ‘realisation’ field then allows one to find examples of syncretisms based on the same form.

Toward the bottom left of figure 3 is a table which contains fields of word classes (nouns, adjectives, verbs etc.). This table can construct sets of word classes to be used in the Domain table. This then enables one to see which word classes typically group together in syncretism domains. The Word
Class table has a relationship with the Domain table, which combines three fields: word class, syntax, semantics. Hence, the relationship between the Word Class table and the Domain table is one-to-many, as a single set of word classes could occur under different syntactic or semantic restrictions.

The table in the top left contains information about languages. There are three fields: Language, Family, Report. The Report field is a hyperlink field which contains a detailed report on the status of syncretism in the inflectional morphology of the language in question. To its right the Language table has the LanguageSyncretismDomain table (LSD table), which is the heart of the database. It combines the information about languages with that about syncretisms and domains. The relationship between the Language table and the LSD table is one-to-many. The relationship between the Syncretism table and the LSD table is one-to-many, and the relationship between the Domain table and the LSD table is also one-to-many. The LSD table also contains a hyperlink field which links to a document containing illustrative paradigms for each syncretism in the database. This thereby allows for quality control by the user, who can see how the analysis, encoded by the choice of morphosyntactic combinations, is arrived at. This is a substantial enhancement, one not often found with typological databases.

3.3.2 Verbal person database
The second database (not included in the original proposal) was an opportunity to explore a different approach to typological databases. It is simpler, reflecting its emphasis on scope over detail. Because the range of both features and values is so restricted, it was possible to use a single, atomic value (e.g. ‘first person singular’ is encoded simply as ‘1SG’, not as separate values for person and number); this allowed us to circumvent the Combination table used in the main database. Thus, each syncretism is encoded as a string of atomic values. It also records properties which might be conditioning factors: tense/aspect/mood, inflectional class, gender of subject, and syntactic context. As with the first database, each instance of syncretism is hyperlinked to an illustrative paradigm.

4 Results
4.1 Databases
We compiled a substantial amount of material for a cross-linguistic typology of syncretism, which comprises the basis for the two relational databases discussed above. Their coverage extends considerably beyond what was originally proposed. The first database records all instances of inflectional homophony for all the inflected word classes in thirty genetically diverse languages (see table 1), comprising 1256 separate entries. The second database covers a more limited phenomenon over a much broader range of languages, cataloguing syncretism of subject person in intransitive verbs in 111 languages. Person marking was chosen for this large-scale survey because, of all inflectional features, it is the one which is most comparable cross-linguistically, the same core values being nearly universally present. Verb forms which mark the person of the object are not included, as that is a typologically distinct phenomenon (see §4.2.2.2).

4.2 A general theory of syncretism
4.2.1 Constraints on syncretism
The goal of much previous work on syncretism has been to model in such a way as to define possible and impossible patterns. As prominent and representative claims we take (i) Carstairs’s (1987) Systematic Homonymy Claim; (ii) Williams’ (1994) ban against contrary syncretic structure; and (iii) Stump’s (2001) Feature Ranking Hierarchy. While these proposals are provocative, they have been based on a highly restricted corpus of examples. Our research has enabled us to evaluate them on a firmer empirical basis. None of the claims is fully supported.

The most significant aspect of each of these models is the way in which morphological stipulation of syncretism is treated. In Williams’ model it is not permitted at all, while Carstairs-McCarthy and Stump attempt to constrain it. Williams proposes that all syncretism is underspecification and is constrained by underlying feature structure. For example, if a language has the feature values a, b, c and d, with the feature structure as shown in figure 4, then it can display syncretism of a + b or c + d, but not of a + c, a + d, b + c or b + d.
However, this model fails when faced with overlapping syncretic patterns, which are frequent cross-linguistically. For example, in Somali verbs (figure 5), 2SG is identical to 3SG feminine in main clauses, but to 3SG masculine in relative clauses. Since each pattern implies a distinct structure, Williams’ model predicts this should not occur. On our analysis, such examples demonstrate the need to admit the overt morphological stipulation of syncretism. This is equally problematic for Noyer’s (1997) notion of Impoverishment, which likewise admits only underspecification, though with a more elaborate feature hierarchy.

<table>
<thead>
<tr>
<th></th>
<th>main clause</th>
<th>relative clause</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>idhi</td>
<td>idhi</td>
</tr>
<tr>
<td>2SG</td>
<td>tidhi</td>
<td>yidhi</td>
</tr>
<tr>
<td>3SG MASC</td>
<td>yidhi</td>
<td>yidhi</td>
</tr>
<tr>
<td>3SG FEM</td>
<td>tidhi</td>
<td>tidhi</td>
</tr>
</tbody>
</table>

‘say (past simple)’

On the other hand, while Carstairs-McCarthy and Stump admit morphological stipulation, their proposed constraints are too restrictive, though they do capture general tendencies. Carstairs-McCarthy’s proposal is, in effect, that in languages where features are separately expressed by distinct morphemes, a referral (essentially, syncretism which is the result of a paradigm-internal borrowing) can only work in the context of a separately expressed feature which is less semantically relevant. For example, number is semantically more relevant than case, so we should expect referral between different number values in the context of a particular case, but not referral between different case values in the context of a particular number. This constraint is violated by phenomena such as case syncretism in Classical Arabic (figure 6): number and case are expressed separately, but the accusative appears to take the form of the genitive in the plural. (Nevertheless, generalisations can still be made about the conditioning role of feature contexts; see §4.3.)

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>hayawa:n-u</td>
<td>hayawa:n-at-u</td>
</tr>
<tr>
<td>GEN</td>
<td>hayawa:n-1</td>
<td>hayawa:n-at-1</td>
</tr>
<tr>
<td>ACC</td>
<td>hayawa:n-a</td>
<td>hayawa:n-at-1</td>
</tr>
</tbody>
</table>

‘animal’

Stump proposes that, if you find in a language a morphologically stipulated syncretism between values of feature \( x \) in the context of feature \( y \), you will not also find the morphologically stipulated syncretism of a value of feature \( y \) in the context of feature \( x \). This constraint is violated by Slovene case and number syncretism: number is syncretic in the context of case (dual + plural are syncretic in the genitive and locative), and case is syncretic in the context of number (there are many patterns of case syncretism specific to the singular, the dual or the plural). All of these must be, by Stump’s criteria, morphologically stipulated syncretisms.

We must conclude that previous attempts to constrain syncretism cannot account for the complete range material which we have investigated. We believe the most promising approach is to construct a nuanced typology of possible formal interpretations, with different limitations applying to each one.

4.2.2 Formal typology
For all the possible differences in implementation, there are in essence only three ways to characterise a syncretic form:

1. Form = function; that is, the syncretic form directly reflects the internal organisation of morphosyntactic features. That is, the syncretic form has a single, semantically coherent meaning. For example, in Burushaski, male gender and female gender, used only with human referents, are syncretised in the plural. Clearly, the syncretic form represents 'human', the union of the two values it syncretises.

2. Underspecification; that the syncretic form is an elsewhere form, which derives its function as a consequence of paradigm structure. For example, in the Nilo-Saharan language Me’en, 2SG and 3SG are syncretic in the past tense. However, the distribution of forms (figure 7) suggests rather that the primary distinction in the past tense is simply between singular lim(u)a and plural limd(u)a; the 1SG formed from the default singular form plus the prefix ki-. Thus, although the meaning of a form such as lim(u)a could be characterised as 'singular past', its function is limited to 2nd and 3rd person, because there is a more specific form to serve for 1st person.

3. Morphological stipulation. The functions that the syncretic form performs have no obvious connection to each other, and must be overtly defined. For example, in some North Sami dialects the comitative singular and inessive-ellative plural function are regularly syncretic throughout the nominal system (Hansson 1996). This cannot be attributed to morphophonological accident, much less to any kind of semantic unity. Only morphology can link the two.

Clearly, no single set of constraints can reasonably apply over these three distinct phenomena. We propose a limited set of constraints that can be imposed on each one.

4.2.2.1 Form = function
If one believes that a syncretic form represents a discrete semantic or syntactic unit, then ideally these units should be recoverable on the basis of morphology-independent criteria. For example, the syncretism of syncretism of male and female genders in Burushaski (§4.2.2) is correlated with a bona fide natural semantic class (human beings). The syncretism of subject and object cases, the most frequent type of case syncretism cross-linguistically (see below), is correlated with the syntactic structure of transitive predicates. Failure to adhere to this criterion may lead to a proliferation of dangerously abstract covert semantic categories, with little in the way of explanatory or predictive power. We have found some cross-linguistically robust patterns which we take as evidence of an underlying semantic or syntactic unity, which are laid out below.

4.2.2.1.1 Case
The multiple intersecting patterns of case syncretism as found in familiar Indo-European languages have constituted the most discussed type of syncretism, precisely because it has been believed, since Jakobson (1936) that these patterns can be taken as a direct reflection of the semantic network which underlies case systems. We looked extensively at case syncretism outside Indo-European and found little support for this view; instead, syncretism correlates with structural syntactic features. Outside Indo-European, case syncretism overwhelmingly involves the core cases used to mark subjects and objects (nominative, accusative, ergative). Most typically there is a split along core/ non-core lines, with the core cases syncretic with each other, or-- much less often-- collapse of all non-core cases with each other. This suggests the legitimacy of a (syntactic) category of core or direct cases which licenses the syncretism. The other common pattern occurs where the core cases are kept distinct from each other, but the marked core case (accusative or ergative) is syncretic with one of
the non-core cases; for example, in Finnish, the genitive case form serves as the accusative for singular nouns, and in Koryak, the locative case form serves as the ergative for second declension nouns. For at least some of the examples where we have historical evidence, this appears to result from the cooption of a non-core case form to establish a distinction between the core cases where none existed previously. There is however wide variation in the choice of non-core case: for example, Russian uses the genitive form for the animate accusative, while East Armenian, with an almost identical case system, uses the dative Ergative case forms may be identical to adnominal cases, to the instrumental, or to the locative.

4.2.2.2 Person
Patterns of person marking syncretism often lead to speculation on the internal composition of person values. Our research shows that there are some strong patterns, which are correlated with number. In the non-singular, or where number marking is independent of person marking, both 1 + 2 and 2 + 3 are common patterns. Since this grouping is also found in free pronouns, this suggests to us that there may be, or have been, some semantic basis to this syncretism. In the singular 2 + 3 overwhelmingly predominates. This appears to be a purely morphological phenomenon resulting from underspecification, since it is typically limited to individual morphological subparadigms, and the syncretic form is frequently the bare stem.

4.2.2.3 Gender
Cross-linguistic generalisations about the semantic licensing of gender syncretism are made difficult by the extreme variety that gender systems display. One prediction, though, appears to hold true: the semantic motivation for gender syncretism in non-singular numbers will be equal to or greater than the semantic motivation in the singular. For example, in Telugu, feminine gender is syncretic with neuter in the singular, while in the plural it is syncretic with the masculine, yielding there a simple human/non-human distinction.

4.2.2.4 Number
The notion of syncretism of particular values of the feature ‘number’ (as opposed to the simple non-distinction of number) arises only in systems where three or more numbers are distinguished (at least singular, dual and plural). In the overwhelming majority of cases one finds dual and plural collapsed, which supports the notion of ‘non-singular’ as a viable semantic unit. However, syncretism of singular and dual is also found; since this may be correlated with semantic features such as animacy, this must be considered as a potentially viable semantic class as well. Syncretism of singular and plural, as opposed to dual, occurs as well, but under such conditions as to suggest it is a result of underspecification, with only the dual being overtly marked.

4.2.2.5 Underspecification
As a basic principle, we consider that there must be some solid justification for regarding the ‘elsewhere’ form as a morphological default. The evidence may vary: e.g. it is a bare stem, or if not, nevertheless displays a tendency to be used beyond its canonical range. Otherwise, underspecification risks being merely a hollow technical device, without explanatory value.

The obvious move taken by many researchers is then to establish a correlation between morphosyntactic markedness and morphological markedness. This is a principle of Natural Morphology, and has been incorporated into some morphological models (Carstairs-McCarthy 1998, Wunderlich and Fabri 1995), by prohibiting morphological rules to make reference to the unmarked values of features. If tenable, such a principle would provide at least a minimal constraint on what can be described by underspecification. However, the notion of ‘markedness’ remains too elusive to be of practical use in declaring possible and impossible patterns. Compare the person markers shown in figure 80 only one singular person is overtly marked, while the other two are syncretically expressed by the bare stem. Every logical possibility is attested, suggesting either that markedness is language-dependent, or that it plays no decisive role in constraining the possibilities. We conclude that the correlation between underlying and surface markedness, while undoubtedly in evidence (witness the fact that nominative and absolutive cases are frequently morphologically unmarked), cannot be encoded as an absolute rule in morphology.
<table>
<thead>
<tr>
<th></th>
<th>Atakapa</th>
<th>Kashmiri</th>
<th>Nez Perce</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>[stem] -o:</td>
<td>[stem]</td>
<td>[stem]</td>
</tr>
<tr>
<td>2SG</td>
<td>[stem]</td>
<td>[stem] -Vth</td>
<td>[stem]</td>
</tr>
<tr>
<td>3SG</td>
<td>[stem]</td>
<td>[stem]</td>
<td>hi- [stem]</td>
</tr>
</tbody>
</table>

There is however one area where certain principles can be laid down, namely with multiple argument marking, where both subject and object are marked on the verb. It is often the case that, given certain combinations of subject and object person, some or all of the information about one of the arguments is omitted. Typically, it is the subject person which is omitted when the object is first or second person, a phenomenon which has been attributed to ‘pragmatic avoidance’ (Heath 1991); high animacy arguments are disfavoured as objects, especially where the subject is also of high animacy. However, even here there may be an element of purely morphological stipulation. For example, in the Tibeto-Burman language Hayu, omission of argument specification partly follows the familiar semantic hierarchy 1st person > 2nd person > 3rd person, but 3PL arguments are marked in preference to 2SG, which must be stipulated.

4.2.2.3 Morphological stipulation
Possible patterns of morphological stipulation are dictated by restrictions both on by diachronic pathways and on formal modelling.

4.2.2.3.1 Diachrony
We identify two primary paths:

(i) Paradigm-internal borrowing. One example is the use of non-core cases in the function of the marked core case, as discussed in §4.2.2.1.1. Another is the substitution of third person markers for second person markers, ultimately with a pragmatic motivation, as found in transitive verbs in Dalabon (Evans, Brown and Corbett 2001).

(ii) Analogical extension of accidental homophonies. For example, in Old Icelandic, 2SG and 3SG indicative present forms fell together in one class of verbs as a result of a general phonological process. This identity was later extended to other verb classes not affected by that phonological change.

4.2.2.3.2 Formal modelling
In the examples just cited, the syncretic form has a clear source, either synchronically or diachronically: the form which serves function x is extended to functions x + y This notion can be extended to the model of morphological structure in the form of directional rules (referrals). However, the status of directional rules in morphology remains controversial, and they are rejected by many in favour of non-directional stipulation (indexing or disjunction). We have shown that directional rules are needed to account for overlapping syncretic patterns, such as found in Somali verbs (see §4.2.1). A non-directional account of such a patterns will run into irresolvable rule conflict; thus if tidhi is the 2SG + 3SG FEM form and yidhi is the 2SG + 3SG MASC form, there is no ready way to determine which form to use for the expression of 2SG. Neither rule is more specific. This difficulty is resolved if we say that the 2SG ‘borrows’ the form either of the 3SG FEM or the 3SG MASC.

4.3 Feature interaction
Words are often inflected for more than one morphosyntactic feature, and some robust generalizations about nominal word classes can be made about the interaction between features whose values are syncretised and those values remain constant (the context). Thus, if case is a context for some other syncretism in language A, then case must be syncretised somewhere in language A; if gender is a context for some other syncretism in language A, then gender must be syncretised somewhere in language A.
5  Activities
In 2001 we were invited to become a partner in the European Thematic Network ‘Language Typology Resource Centre’ run from the University of Utrecht. We made a presentation on our Syncretism Database at Utrecht and the organisers were particularly excited by it. In the original proposal we promised to make the database available over the web. Because of the interest of the thematic network and additional funding they made available we have been able to create an enhanced web interface which is straightforward to use. The first ‘live’ presentation of the database was 26.09.02 and it has been continuously available since then. The European ‘Language Typology Resource Centre’ site in Utrecht will have links to our database and we have independently submitted metadata directly to the Open Languages Archives Community at the University of Pennsylvania. We have also advertised the database on the Linguist List.

6  Outputs
6.1 Online resources
Both databases are freely available online to all interested researchers at <http://www.smg.surrey.ac.uk>. We have created a simple, user-friendly interface in order to encourage maximum use of the resource. In addition, through our participation in the Typological Database Project and the Open Language Archives Community (see §5), the databases will form part of a larger international network.

6.2 Annotated bibliography of syncretism
The annotated bibliography contains 100 entries, including works in German, Russian, Polish and French, and is available at <http://www.surrey.ac.uk/LIS/MB/Bibliography.htm>.

6.3 Presentations
We have given thirty-one presentations in Britain, the US, Germany, Russia, Sweden, the Netherlands, Austria, Switzerland and Malta. Four of these have since been published.

6.4 Publications
Six papers have been published so far. In addition, two chapters in the World Atlas of Language Structures have been completed and are forthcoming from Oxford University Press (for information, see <http://www.eva.mpg.de/lingua/wals/>), and a monograph on syncretism is in preparation (see §8).

7  Impacts
7.1 Databases and online resources
We have had a number of positive responses to the preliminary versions of the databases and annotated bibliography that were placed online. Through our participation in workshops organised under the auspices of the Typological Database Project in the Netherlands, we have been able to make a substantial contribution to the design of that collaborative database network.

7.2 Theoretical linguistics
The presentations and publications listed in §6 show that a broad spectrum of linguists is already familiar with the project. There is a lively discussion of some of our work in two of the articles in the most recent Yearbook of Morphology (2001 edition).

7.3 Linguistic typology
Our two chapters in the World Atlas of Language Structures (see §6.4) are an important contribution, as they are among the few (ten out of a total of 133) to focus on morphological issues.

8  Future Research Priorities
In 2001 we concluded a contract with Cambridge University Press for a book entitled Inflectional Syncretism, to be part of the Blue Series (Cambridge Studies in Linguistics). This will be an opportunity to consolidate and synthesise the findings of the project, as well as expand the coverage in certain areas -- especially by further exploring the impact of diachrony on the model we have developed.
References
Appendix: Outputs

Electronic Resources

The following articles and presentations have arisen wholly or in part from the project:

Publications

Presentations
Baerman, Matthew. 2000c. Is there a semantic basis for person syncretism in verb forms? ESRC seminar series 'Methods and Models in Morphology', University of Sussex, November 17.
Baerman, Matthew. 2001c. Resolving rule competition. ESRC seminar series 'Methods and Models in Morphology', University of Essex, December 7.


Brown, Dunstan. 2001a. Features as context and content of syncretism ESRC seminar series 'Methods and Models in Morphology', University of Surrey, February 9.


Corbett, Greville. 2000g. Possible words: four inflectional phenomena for defining the extremes. Falmer Language Group Seminar, University of Sussex, May 15.


Corbett, Greville. 2000g. Domains of syncretism: a typological perspective on Slavonic. 26th Meeting of the Internation Commission on the Grammatical Structure of the Slavonic Languages, Männedorf, near Zürich, October 22-23.


12