Suppletion: frequency, categories and distribution of stems

Andrew Hippisley, Marina Chumakina, Greville G. Corbett and Dunstan Brown[*]

University of Surrey

Andrew Hippisley
Department of Computing
School of Electronic and Physical Sciences
University of Surrey
Guildford, Surrey
GU2 7XH
UK

a.hippisley@surrey.ac.uk

Abstract
Suppletion is where the word-forms of the same lexeme have phonologically distinct stems. A study of thirty languages shows it to be surprisingly widespread, suggesting resistance to the pressure of paradigmatic levelling. While a major factor in its preservation appears to be the high frequency of the items that display it, two other factors are in operation, the type of inflectional category involved and the nature of the distribution of stems.

1   Introduction

Suppletion, where the set of morphosyntactic forms of a lexeme contains phonologically unrelated stems, is surprisingly pervasive. Although in a given language it typically involves a very small proportion of the lexicon, in most cases the items that display it have unusually high token frequencies. Moreover, based on a survey of thirty unrelated or distantly related languages, suppletion was found to occur in the overwhelming majority. As such, suppletion impacts on theories of the lexicon, language change and language acquisition, and for this reason "is of considerable interest" (Mel´čuk 2000: 511). The usual observation about suppletion is that it affects high frequency items. We wish to extend this characterisation of suppletion by identifying two additional properties. We shall show how this yields a more robust account that may help to explain why suppletion is maintained in the lexicons of so many languages, even in those cases where there is clearly no frequency effect. In addition to frequency (property A), property B involves the inflectional categories that provide the suppletion context, and property C concerns the paradigmatic distribution of the stems. Our argument is that these three properties combine in a particular way to preserve suppletion in the lexicon. Cases of suppletion apparently uncharacteristic for one property can be accounted for by showing them to be entirely characteristic for the two other properties.

In section 2 we give a working definition of suppletion. Section 3 is an overview of our suppletion database, a repository of analyses of thiry languages. In section 4 we present the first of the three properties identified with suppletion, high token frequency. The second property, discussed in section 5, concerns the inflectional categories that participate in suppletion. We argue that suppletion is typically restricted to the class of 'inherent' categories such as number marking on nouns. The third and final property is more complex as it is more abstract and concerns the nature of the distribution of the phonologically distinct stems amongst the morphosyntactic cells within a paradigm. This is discussed in section 6. All three properties are drawn together in section 7, where we show how exceptions to one property show typical behaviour for the remaining two properties. Finally, in section 8 we present one example that challenges our claims that a suppletive lexeme displays a combination of two or all of properties A, B and C. In a range of languages the copula is suppletive and has a high frequency, thereby adhering to Property A; at the same time it is found to display suppletion in contextual categories, thereby flouting Property B, and have its own unique morphological patterning, thereby flouting Property C.

2   Defining suppletion

Before discussing the three properties that are typically found with suppletion, we require

a working definition of suppletion, and for this we draw on Mel´čuk's (1994: 343) definition in (1).

(1)   "Suppletion is a relation between signs X and Y such that the semantic difference...between X and Y is maximally regular...while the phonological difference is maximally irregular."

For example, in Russian, stems are inflected to mark number: *student* : *student-i* 'student : students' where the presence of a suffix distinguishes the plural from the singular form. Apart from this, both forms are identical. This is the typical situation. However, we also find instances of suppletion, for example *reb´on(o)k* : *det´-i* 'child : children', where the stems of the word-forms are distinct. In terms of Mel´čuk's definition, along the dimension SINGULAR : PLURAL the relationship between the signs X and Y is semantically maximally regular (both have the sense of 'child') but phonologically maximally irregular: there are no possible phonological rules in Russian that can map the stem *reb´on(o)k* to the stem *det´-*. On the other hand, *student* : *student-i* 'student : students' is not a case of suppletion since the relationship between the two signs is maximally regular both semantically and phonologically.

The definition of suppletion above can have either a broad or a narrow interpretation, and throughout we assume the narrow one. First, one may view some cases of derivation as examples of suppletion, and indeed such cases are presented by Mel´čuk. Our claim, however, revolves entirely around inflectional suppletion. Second, the narrow interpretation includes phonologically unrelated stems, and excludes affixes. Again, we follow the narrow interpretation. Finally, some view suppletion in terms of degree (see Carstairs-McCarthy 1994: 4411), where 'weak' suppletion simply means that the stems do not have full identity. Under this definition, English *child : children* would be an example of suppletion. Again, we use the narrow interpretation, namely that phonologically the relationship between stems is <u>maximally</u> irregular. Finally, we take a synchronic view of suppletion; stems which meet the definition are treated as suppletive irrespective of their etymology.

3   The Surrey Suppletion Database
The idea behind creating the Surrey Suppletion Database was to provide easily accessible information about suppletion to facilitate typological claims about the phenomenon. To ensure a broad range of examples we followed a number of criteria in selecting the languages.   First, we ensured genetic and areal diversity. Second, given that our interest is in suppletion, a language must have inflectional morphology. Finally, a practical consideration was that, where possible, for a given language there should exist a good grammar, dictionary, and a specialist whom we could consult when necessary.   The table gives the 30 languages that have been analysed to date with their family affiliation.

Table 1: Languages and their families in the Surrey Suppletion Database

Archi	Nakh-Daghestanian
Basque	Basque
Chichewa	Niger-Congo
Georgian	Kartvelian
Guaraní (Paraguayan)	Tupí
Hebrew	Semitic
Hua	Trans-New Guinea
Hungarian	Uralic
Itelmen	Chukotko-Kamchatkan
Kannada	Dravidian
Kayardild	Australian, Tangkic
Ket	Yenisei Ostyak
Koasati	Muskogean
Komi	Uralic
Limbu	Sino-Tibetan
Mayali	Australian, Gunwingguan
Maybrat	West Papuan
Navajo	Athabaskan
Ojibwa	Algic
Palauan	Austronesian
Qafar	Cushitic
Tetelcingo Nahuatl	Uto-Aztecan
Tarma Quechua	Quechuan
Totonac	Totonacan
Turkana	Nilo-Saharan
Xakass	Turkic
Yacaltec	Mayan
Yimas	Sepik-Ramu
Yukaghir	Yukaghir
Yup'ik	Eskimo-Aleut

In order to support consistency in data entry we have created a relational database, which allows us to enter and store information without redundancy. By treating the data in terms of a number of tables with relationships between them, we are able to place constraints on the information entered. Figure 1 gives the underlying structure of the database.
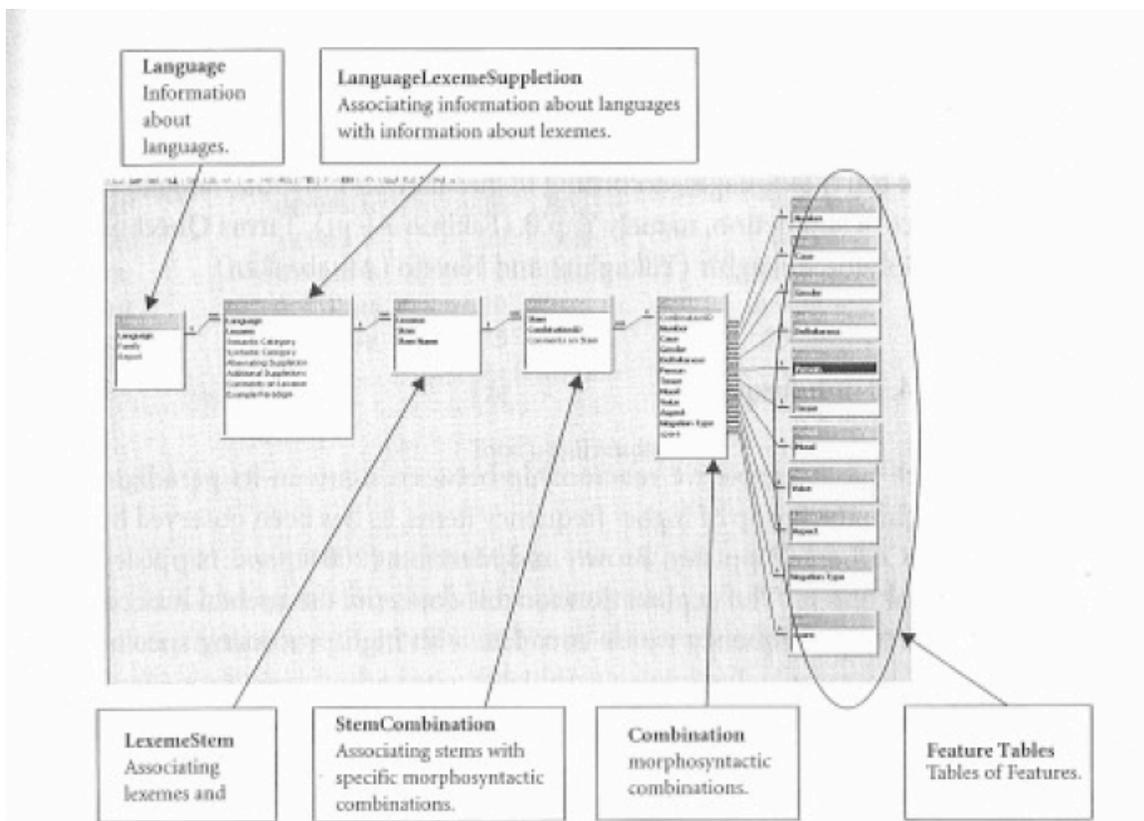
Figure 1. The Surrey Suppletion Database

To the right of figure 1 we see ten tables of features (`Number, Case,` etc.) plus an additional 'spare' table. These eleven tables are encircled in the figure. The values from these tables can combine in a morphosyntactic combination, as represented by the `Combination` table. As a value can occur in more than one morphosyntactic combination, the relationship between feature values and morphosyntactic combinations is one-to-many. We can use the morphosyntactic combinations to define the morphosyntax associated with particular stems. In the `StemCombination` table, stems are associated with the morphosyntactic content of which they are the expression in form. Stems are language-specific items, whereas morphosyntactic combinations are not. The fact that morphosyntactic combinations generalize across languages means that one morphosyntactic combination in the database may be associated with many stems (each belonging to a different language). Hence the relationship between the `Combination` table and the `StemCombination` table is one-to-many. The `LexemeStem` table provides a triple of information: the lexeme name (lexeme being an abstraction over a whole paradigm); a stem name; a description of the stem. The relationship between the `LexemeStem` table and the `StemCombination` table is one-to-many, as a language specific stem is constrained to be described once, but this stem could in principle occur in more than one morphosyntactic combination (in the `StemCombination` table). The beauty of having a separate `LexemeStem` table and `StemCombination` table is that we can then describe stems both in terms of the morphosyntax and in terms of their arbitrary morphological function (in the field 'stem name' in the `LexemeStem` table). This allows us to analyse a stem's morphosyntactic and 'morphomic' properties (Aronoff 1994). The

table `LanguageLexemeSuppletion` brings together the information about instances of suppletion and languages, and introduces further fields, such as semantic categories (the lexical semantics of the items involved), the syntactic categories (word classes), whether alternative stems can be involved in the same suppletion relationship, and if the lexeme has additional instances of suppletion. There is also a hyperlink to example paradigms. The language table provides information about the languages in the table, together with an individual report on each language. It is interesting to note that out of thirty languages, according to our analysis only four languages have no instance of suppletion, namely Yup'ik (Eskimo-Aleut), Tarma Quechua (Quechuan), Kolyma Yukaghir (Yukaghir) and Navajo (Athabaskan).

4   Property A: High frequency

A lexeme which has a suppletive relationship between stems in its paradigm usually falls within the group of higher frequency items, as has been observed by Bybee (1995), Corbett, Hippisley, Brown and Marriott (2001), and Hippisley (2001), amongst others. The explanation for this concerns the mental lexicon and its storage. High frequency words correlate with high processing speeds, evidenced by, for example, lexical decision times, suggesting a memory effect: the more frequently an item is accessed, the stronger the memory traces, hence the shorter the response times (see for example Clahsen, Eisenbeis, Hadler and Sonnenstuhl 2001, and the discussion in Hippisley 2001). The implication is that high frequency items, that is 'salient' items, are stored differently. As an example, we consider the token frequencies of the suppleting lexeme for 'child' *reb´on(o)k : det´-i* in Russian in Table 2 and the non-suppleting lexeme for 'girl, young woman' *devušk-a : devušk-i* in Table 3. The frequencies are taken from a one million word corpus of Russian, the Uppsala Corpus.

Table 2: Russian *reb´onok : det´-i* 'child', from the Uppsala corpus

	singular			*plural*	
Nom	reb´onok	44	Nom	det´-i	175
Acc	reb´onk-a	20	Acc	det-ej	99
Gen	reb´onk-a	70	Gen	det-ej	134
Dat	reb´onk-u	8	Dat	det´-am	40
Inst	reb´onk-om	15	Inst	det´-m´i	34
Loc	reb´onk-e	4	Loc	det´-ax	6
total singular	161	total plural	488		
Total occurrences: 649					

Table 3: Russian *devušk-a : devušk-i* 'girl', from the Uppsala corpus

	singular			plural	
Nom	devušk-a	57	Nom	devušk-i	25
Acc	devušk-u	18	Acc	devušek	7
Gen	devušk-i	22	Gen	devušek	15
Dat	devušk-e	8	Dat	devušk-am	3
Inst	devušk-oj	20	Inst	devušk-am´i	6
Loc	devušk-e	3	Loc	devušk-ax	1
total singular	128		total plural	57	
Total occurrences: 185					

If we compare suppletive pair *reb´on(o)k : det´-i* with non-suppletive *devušk-a : devušk-i* for <u>absolute </u>frequency we see that the suppletive item has a greater frequency: the total number of occurrences in the case of *reb´on(o)k : det´-i* is 649 and in the case of *devušk-a : devušk-i* is the much lower figure of 185. Another way of measuring the relationship between suppletion and high frequency is to look at the frequency distribution within the paradigm of a single lexeme and compare that with other lexemes. In some cases the distribution of phonologically unrelated stems is according to the singular and plural subparadigms, as in our example in Table 2. We take the <u>relative </u>frequency to be the proportion of the full set of occurrences of a lexeme represented by plural occurrences. In Table 2 we see that for the suppleting item the proportion is 75%, i.e. of the full set of 649 occurrences, 488 are plural occurrences. The <u>relative</u> frequency is again much lower for the non-suppleting item which is around 31% and in fact close to the typical <u>relative</u> frequency in the corpus. Thus whether dealing with proportions or absolute numbers, there is good evidence that suppletion is related to high frequency.

We assume that there is a correspondence between a lexical item's frequency and its salience, where high frequency corresponds to high salience. This is the standard assumption in many text analysis techniques, such as those found in information retrieval (Sparck Jones 1999: 261) To support Property A, in some instances we have direct evidence of frequency. In others we take salience as indirect evidence, assuming that certain items (like 'go', 'mother', 'child') are salient across languages. It is then not surprising that the same semantic class is represented by suppleting lexical items in a number of different languages (see Table 4).

Table 4: Lexical salience and suppletion

	Semantic class	Language	Category	Suppletion
a.	'child'	Russian	NUMBER	*rebʹonok* (SG) / *detʹ-i* (PL)
b.		Xakass		*pala* (SG) / *olɣan-nar* (PL)
c.		Komi		*kaga* (SG) / *čelʹadʹ* (PL)
d.		Turkana		*i-kɔ́ku* (SG) / *Ni-deʕ* (PL)
e.	'this'	Hebrew		*ze* (M.SG) / *ele* (M.PL)
f.	's/he', 'it'	Turkana		*ŋèsiʕ* (SG) / *kèciʕ* (PL)
g.		Itelmen		*enna* (SG) / *itX* (PL)
h.	'say'	Turkana	MOOD	*è-bàl-aʕ* (IND.PAST.3.SG) / *tɔ-maʕ* (SUBSEC.3.SG)
i.		Hebrew	MOOD	*l-emor* (IND.INF) / *tagid* (IMPERATIVE)
j.		Georgian	TENSE	*amb-obs* (PRES.3.SG) / *i-tʼqʼv-is* (FUT.3.SG)
k.	'come'	Hungarian	MOOD	*jön-ni* (IND.INF) / *gyere* (IMPERATIVE)

Exceptions to this first property (high frequency) are found in Archi, a Nakh-Daghestanian, Lezgian language, and Hua, a Trans-New Guinea language. In Archi there is a set of suffixes that are used to mark plural number for nouns, two of which are shown in examples (2) and (3), from Kibrik(1977: 30):

> (2) 'price' *baha* (sg) / *baha-ttu* (pl)

> (3) 'place' $biq\text{'}_o$ (sg) / $biq\text{'}_o\text{-}mul$ (pl)

The word for 'corner of a sack' (Kibrik 1977: 46), which we assume to have relatively low salience, displays suppletion as shown in (4), and therefore represents a counterexample to Property A.

> (4) 'corner of a sack' *bičʹni* (sg) / *boždo* (pl)

In Hua, polarity is marked by the prefix *ʻa-*, as shown in (5) for the verb *mie* 'give' (Haiman 1980: 193):

> (5) 'give' *mie* (affirmative) / *ʻa-mie* (negative)

The prefix provides a context for regular alternation in a sub-type of verbal compounds consisting of an adjectival argument and a form of the verb *hu* 'do', which in such constructions functions as a copula. The polarity marker is prefixed to the copula. In such cases copula *hu* alternates with *fu* (Haiman 1980: 194) . This is shown in (6).

> (6) 'be good' *soko hu* (affirmative) / *soko ʻa-fu* (negative)

The verbal compound *kta hu* can be glossed as 'be heavy', and is (presumably) a low salience lexical item. When negated with the prefix *'a-*, the *hu* stem alternates with the phonologically unrelated *pri* stem, and as such provides another counterexample to Property A (Haiman 1980: 199):

(7) 'be heavy' *kta hu* (affirmative) / *kta 'a-pri* (negative)

5   Property B: inherent inflection

The first property we associated with suppletion, high frequency, identifies a particular class of lexical items. The second property, Property B, identifies the inflectional categories that provide the context for suppletion. Booij (1996) distinguishes two types of inflectional category based on their interaction with the syntactic context. In <u>contextual</u> inflection the presence and value of the category is dictated by the syntax, for example agreement markers on targets. In <u>inherent</u> inflection, on the other hand, the syntactic context does not play any role in the selection of a category and its value, for example, number marking on a noun controller is not conditioned by any other constituent in the clause. The partitioning of inflection into these two classes is diagrammed in Figure 2 which has examples of each class.
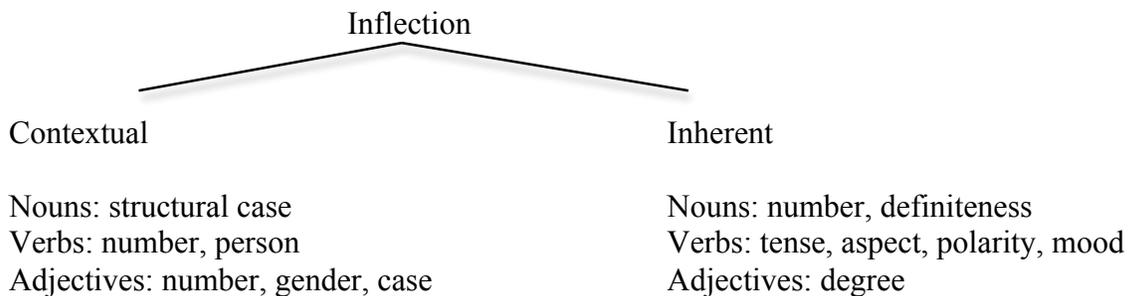
Inflection

Contextual   Inherent

Nouns: structural case   Nouns: number, definiteness
Verbs: number, person   Verbs: tense, aspect, polarity, mood
Adjectives: number, gender, case   Adjectives: degree

Figure 2. Contextual and inherent inflection.

From Figure 2 we see that a morphosyntactic property such as Number can be contextual and inherent. Number marking on the noun is an inherent inflectional category; since verbs are typically targets in agreement, in those cases number marking on verbs is contextual inflection. Given this division amongst inflectional categories, we claim that it is the inherent inflectional categories that typically provide the context for suppletion (Corbett 1999). In Table 4, (a) to (g) are examples of number for nouns, an inherent category; and (h) to (k) show inherent categories for verbs, namely mood and tense.

        Related to Booij's inherent inflection is what Bybee calls <u>relevance</u> of morphological category. A category is more relevant if it more drastically affects the semantics of the stem. For example, for the semantic content of a verb aspect has a more drastic effect than person or number (Bybee 1985: 57). Her 'relevant' categories coincide largely with Booij's inherent categories. Interestingly, our property B is similar to her point that it is within these categories we find suppletion, and she offers an explanation in keeping with her architecture of the mental lexicon. Morphologically related forms are bound together in networks through semantic connections. The more

relevant the morphological category, the weaker the connection, and this is reflected by a weaker phonological relationship. If suppletion is to occur, it is in this context:

> "weaker semantic connections are reflected in phonological form by a greater degree of stem change, or even suppletion among related forms." Bybee (1995: 429).

Exceptions to Property B will be cases of suppletion which clearly involve contextual inflection rather than inherent. An example is the Danish word for 'small' which is suppletive in two inflectional contexts. The first is Degree, as shown in (8) and (9).

(8)   en   lille   øl
         a   small.positive   beer
         'a small beer'

(9)   en   mindre   øl
         a   small.comp   beer
         'a smaller beer'

This is consistent with Property B since for adjectives Degree is inherent (Figure 2). However, the same lexical item also displays suppletion in Number: as indicated in Figure 2, Number for adjectives is a contextual category, as adjectives are targets in agreement relations that involve number. In (10) and (11) we see the alternation of two phonologically unrelated stems, *lille* and *små*, conditioned by the number of the controller *barn* 'child'.

(10)   et   lille   barn
         a   small.sg   child.sg
         'a small child'

(11)   små   børn
         small.PL   child.PL
         'small children'

This can be compared to the regular adjective *glad* 'happy' where Number is marked through suffixation in (12) and (13).

(12)   en   glad   hund
         a   happy.SG   dog.SG
         'a happy dog'

(13)   glad-e   hund-e
         happy-PL   dog-PL
         'happy dogs'

Another exception to Property B is represented by Bagwalal, a Nakh-Daghestanian language. In the past tense, Bagwalal verbs use suffixes and circumfixes to agree with their absolutive controllers in gender and number. In the singular, three genders are distinguished: masculine by the prefix *w-*, feminine by the prefix *j-* and neuter by the prefix *b-*, as shown in (14) to (16).

(14) *waša*   ***w=iRi***
      boy   M=stop.past
      'the boy stopped'

(15) *jaš*   ***j=iRi***
      girl   F=stop.past
      'the girl stopped'

(16) *ʕama*   ***b=iRi***
      donkey N=stop.past
      'the donkey stopped'

In the plural only two genders are distinguished: human through the circumfix *b-...-r* and non-human using the prefix *r-*; compare (17) and (18).

(17) *waša-bi / jaš-i*   ***b=iRi=r***
     boy-PL / girl-PL   HUM.PL=stop.PAST=HUM.PL
      'the boys/girls stopped'

(18) *ʕama-bi*   ***r=iRi***
     donkey-PL   NONHUM.PL=stop.PAST
     'the donkeys stopped'

The verb 'come' is suppletive, having two phonologically distinct stems in the past tense, **-e:-** and **-a:-** . What is interesting is that the alternation of these stems is dictated by the syntactic context, specifically gender and number agreement. The stem **-e:-** is selected for +Human +Plural agreement, and the stem **-a:-** for all other agreement contexts. This is shown in (19), (20) and (21). Stems are is bold.

(19) *aram-di*   ***b=e:=r***
     person-PL   hum.pl=come.past=hum.pl
     'people came'

(20) *aram*   ***w=a:***
     person   m=come.past
     'a person came'

(21) *zin-a*   ***r=a:***
     cow-PL   nonhum.pl=come.past
     'cows came'

Further counterexamples to Property B involving copulas are discussed in section 8 below.

To sum up, Property B covers the inflectional categories that typically provide the context for suppletion, and we have shown that these are drawn from the inherent subset. While this is a strong tendency we have also identified exceptions to this property, namely cases of suppletive items which play the role of target within an agreement relation.

6   Property C: Morphologically systematic

The final property concerns the distribution of phonologically distinct stems over the morphosyntactic cells within a paradigm. The claim is that the patterning of these stems is not unconstrained but is determined by the morphological system in place. In other words, the distribution of phonologically unrelated stems in cases of suppletion follows the distribution of stems of more regular lexemes. As an example we can consider Latin. Regular Latin verbs are treated traditionally as having a stem inventory, since they display three distinct shapes within the paradigm. Taking the regular first conjugation verbs *am-ō* 'love' and *laud-ō* 'praise', Table 5 shows how the stems are distributed among the morphosyntactic categories Present Active Infinitive, Perfect Active and Perfect Passive Participle.

Table 5: Stems of *am-ō* and *laud-ō*

Present Active Infinitive	Perfect Active	Perfect Passive Participle
amā-re	amāv-i	amāt-us
laudā-re	laudāv-i	laudāt-us

The full range of categories associated with each of the three stems is given in Table 6.

Table 6: Distribution of stems in the Latin verbal paradigm.

am-ō	amāv-ī	amāt-um
stem 1	stem 2	stem 3
Active: - present - future - imperfect Passive: - present - future - imperfect	Active: - perfect - future perfect - pluperfect	Active: - future participle Passive: - perfect - future perfect - pluperfect - perfect participle Supine

Using data of this kind, Aronoff (1994) shows that stem distribution need not be morphosyntactically motivated: a stem may serve for unrelated categories, such as the third stem being used for both Passive and Active voice, and both Future and Perfect tense. The distribution reflects the morphological system in place, and it is the system that we claim governs the stem distribution in cases of suppletion. In the case of Latin suppletion, we expect to find a distribution of phonologically unrelated stems according

to the partitioning of the verbal paradigm in Table 6. The word for 'carry' *fer(ō)* is suppletive and the distribution of its stems in Table 7 shows that this expectation is met.

Table 7. Distribution of stems for Latin suppletive *fer-ō*.

fer-o	*tul-ī*	*lāt-um*
stem 1	stem 2	stem 3
Active: - present - future - imperfect Passive: - present - future - imperfect	Active: - perfect - future perfect - pluperfect	Active: - future participle Passive: - perfect - future perfect - pluperfect - perfect participle Supine

The notion that the pattern of suppletion is constrained by pre-existing structures in the paradigm has been discussed with regard to historical restructuring of the paradigm in Romance. Maiden (2002) discusses a morphological template in Romance that distributes alternant stems in a systematic way. Some verbs have one set of forms for the Present singular and third plural, and another for the rest of the paradigm, i.e. for first and second plural. This is shown for 'sit' in Italian and 'come' in French in Table 8.

Table 8: The lexical items 'come' in French and 'sit' in Italian

French 'come'		*Italian 'sit'*	
SINGULAR		SINGULAR	
1	viens	1	siedo
2	viens	2	siedi
3	vient	3	siede
PLURAL		PLURAL	
1	venons	1	sediamo
2	venez	2	sedete
3	viennent	3	siedono

In cases of suppletion, such as the verb 'go', the template acts to 'regulate' the distribution of the forms, as in Table 9.

Table 9: The lexical item 'go' in French and Italian

French		Italian	
SINGULAR		SINGULAR	
1	vais	1	vado
2	vas	2	vai
3	va	3	vá
PLURAL		PLURAL	
1	allons	1	andiamo
2	allez	2	andate
3	vont	3	vanno

Aski (1995) is a detailed account of the history of suppletion for the lexical item 'go' in Romance, and is entirely based on the notion of morphological templates. While the focus is on the origins of suppletion, she concludes that templates as well as high frequency account for why suppletion is preserved.

> "..the resistance to analogical levelling demonstrated by suppletive forms is due not only to the frequency of the forms…but also to the fact that they adhere to a conjugational pattern that renders them less anomalous." (Aski 1995: 427).

Our expectation is that the morphological system will dictate the pattern of suppletion. A clear counterexample is seen in the Slovene nominal paradigm. Slovene marks singular, dual and plural Number. Where there is more than one stem in a noun's paradigm, the distribution is as in *grâd* 'castle' (Table 10): the plain stem is used for singular, and the augmented stem for dual and plural (Priestly 1993: 400-402), as indicated by the shading.

Table 10: Slovene stem distribution.

Singular	Dual	Plural
grad	grad-ov-a	grad-ov-i

In a case of suppletion, we therefore expect the same distribution amongst the phonologically unrelated stems. However, for *člóvek* 'person' we find something different, as shown in Table 11, where one stem serves for the singular and dual, and the other for the plural.

Table 11: Slovene counterexample to Property C

Singular	Dual	Plural
človek	človek-a	ljudj-e

6.1   Georgian verbs and the morphological system as a constraint on syncretism
Georgian, a Kartvelian language, presents a more complex picture. There are verbs that show suppletion for tense, aspect and mood. Of the six Georgian suppletive verbs noted and discussed in Hewitt (1995: 446-501), most of them display Property C, i.e. the distribution of the phonologically unrelated stems is determined by the morphological system that characterises regular verbs. There is, however, a counterexample to this property. To outline the Georgian situation we will start with minimal preliminary information about Georgian verbal morphology, then present morphological patterns for regular non-suppleting verbs. This will put us in a position to determine whether verbs

that exhibit suppletion adhere to Property C by displaying stem distribution regulated by morphological patterns.

Traditionally tense-aspect-mood forms are divided into three series or 'screeves' and four if we take into account the sub-division of series I (Table 12). Each series is associated with a specific set of affixes that apply to the verb and with a specific case encoding of the verbal arguments (Hewitt 1995: 218).

Table 12: Georgian verbal series.

Series I	Series II	Series III	
Present sub-series	**Future sub-Series**		
Present Indicative	Future Indicative	Aorist Indicative	Perfect
Present Subjunctive	Future Subjunctive	Aorist Subjunctive	Pluperfect
Imperfect Indicative	Conditional		Third Subjunctive

In the structure of a verbal word, eleven morphological slots are recognised, as in (22). Examples of a Georgian verb form are given in (23) for *dagic'eres* 'they wrote it for you' and (24) for *davaparebt* 'we will put a cover over (it)' (roots are bold-faced):

(22)	preverbs-AGR-version-**ROOT**-CAUS-PASS-theme-PF/stative-IPF-mood-AGR

(23)	da-	g-	i-	**c'er-**	es	
	preverb	2SG (Set B)	version	root	3PL (Set A)	

(24)	da-	v-	a-	**par-**	eb-	t
	preverb	1PL (Set A)	version	root	theme	1PL (Set A)

The formants we will concentrate on are the preverb, agreement affixes, root and theme. There are two sets of agreement affixes, traditionally referred to as Set A and Set B, the hyphen showing their position relative to the stem (Table 13).

Table 13: Pronominal agreement affixes in Georgian

	Set A		*Set B*	
	SINGULAR	PLURAL	SINGULAR	PLURAL
1	v-	v- -t	m-	gv-
2	ø/x-	ø/x- -t	g-	g- -t
3	-s/-a/-o	-(a/e)n/-es/-nen	ø/s-/h-	ø/s-/h- (-t)

Table 14 represents the correlation between the sets of agreement affixes and the verbal arguments in different series.

Table 14: Distribution of Georgian agreement affixes

Series I		Series II			Series III	
Subj	DO, IO	Subj	DO	IO	Subj	DO
Set A	Set B	Set A	Set B	Set B	Set B	Set A

Let us consider the conjugation of two regular Georgian verbs, 'build' and 'write'. We will not focus on Person distinctions because for verbs it is a contextual inflectional category, and as discussed in section 5 suppletion typically involves inherent inflection. We therefore give third person singular forms only.

Table 15: Georgian regular verbs

Series I		Series II	Series III
Present sub-series	*Future sub-series*		
Present Indicative a-**šen**-eb-s 'build' **c'er**-s 'write'	Future Indicative a-a-**šen**-eb-s 'build' da-**c'er**-s 'write'	Aorist Indicative a-a-**šen**-a 'build' da-**c'er**-a 'write'	Perfect a-u-**šen**-eb-i-a 'build' da-u-**c'er**-i-a 'write'
Present Subjunctive a-**šen**-eb-de-s 'build' **c'er**-de-s 'write'	Future Subjunctive a-a-**šen**-eb-de-s 'build' da-**c'er**-de-s 'write'	Aorist Subjunctive a-a-**šen**-o-s 'build' da-**c'er**-o-s 'write'	Third Subjunctive a-e-**šen**-eb-in-o-s 'build' da-e-**c'er**-o-s 'write'
Imperfect Indicative a-**šen**-eb-d-a 'build' **c'er**-d-a 'write'	Conditional a-a-**šen**-eb-d-a 'build' da-**c'er**-d-a 'write'		Pluperfect a-e-**šen**-eb-in-a 'build' da-e-**c'er**-a 'write'

Besides the series division which defines a morphosyntactic system, it is worth asking whether or not there are purely morphological patterns for more regular verbs which would determine the distribution of suppletive stems. One such system is a two-way contrast based on the presence of a prefix: all the forms besides those in the Present Sub-series have preverbs - *a*- for 'build' and *da*- for 'write'. The first morphological pattern, what we call Pattern 1, can be established based on this division. This partitioning of the paradigm is shown in Table 16 for the verb 'to choke', where the preverb is *da*-.

Table 16: Morphological Pattern A and the regular verb 'to choke'.

Series I		Series II	Series III
Present sub-series	*Future sub-series*		
Present Indicative	Future Indicative	Aorist Indicative	Perfect
a-**xrč**-ob-s	da-a-**xrč**-ob-s	da-a-**xrč**-o	da-u-**xrčv**-i-a
Present Subjunctive	Future Subjunctive	Aorist Subjunctive	Third Subjunctive
a-**xrč**-ob-de-s	da-a-**xrč**-ob-de-s	da-**a-xrč**-o-s	da-e-**xrč**-o-s
Imperfect Indicative	Conditional		Pluperfect
a-**xrč**-ob-d-a	da-a-**xrč**-ob-d-a		da-e-**xrč**-o

If we turn to suppletion, we have an expectation for an item which has two phonologically unrelated roots, namely that one root will be used for the Present Sub-Series and the other for the rest. And indeed this is exactly what we find for two suppletive verbs 'give' and 'be seated', the latter partly defective as it lacks forms for Present Subjunctive and Imperfect Indicative. Each has two distinct roots: 'give' has *-jl-*

and *-c-* and 'be seated' has ***-zi-*** and ***-d-.*** These are distributed according to Pattern 1 (Table 17).

Table 17: Morphological Pattern A, suppletive verbs 'give': *-jl-/-c-* and 'be seated': zi/-ǰd-

Series I		*Series II*	*Series III*
Present sub-series	*Future sub-series*		
Present Indicative	Future Indicative	Aorist Indicative	Perfect
a-**jl**-ev-s 'give' **zi**-s 'be seated'	mis-**c**-em-s 'give' i-**ǰd**-eb-a 'be seated'	mis-**c**-a 'give' i-**ǰd**-a 'be seated'	miu-**c**-i-a 'give' m-**ǰd**-ar-a 'be seated'
Present Subjunctive	Future Subjunctive	Aorist Subjunctive	Third Subjunctive
a-**jl**-evde-s 'give'	mis-**c**-emde-s 'give' i-**ǰd**-ebode-s 'be seated'	mis-**c**-e-s 'give' i-**ǰd**-e-s 'be seated'	mie-**c**-e-s 'give' m-**ǰd**-ariq'o-s 'be seated'
Imperfect Indicative	Conditional		Pluperfect
a-**jl**-evd-a 'give'	mis-**c**-emd-a 'give' i-**ǰd**-ebod-a 'be seated'		mie-**c**-a 'give' m-**ǰd**-ariq'-o 'be seated'

Another morphological pattern can be established for verbs that change the root vowel. Such verbs make a three-way contrast. While the Pattern 1 contrast is preserved, i.e. the Present Sub-Series is contrasted with everything else through the absence of the preverb, another contrast is made based on the vowel alternation of the root; this separates Series II and III from the rest of the paradigm. Table 18 shows the verb 'to catch' whose stem distribution is determined by morphological Pattern 2.

Table 18: Morphological Pattern B, verb 'catch'

Series I	*Series II*	*Series III*	
Present sub-series	*Future sub-series*		
Present Indicative	Future Indicative	Aorist Indicative	Perfect
i-**č'er**-s	da-i-**č'er**-s	da-i-**č'ir**-a	da-u-**č'er**-i-a
Present Subjunctive	Future Subjunctive	Aorist Subjunctive	Third Subjunctive
i-**č'er**-de-s	da-i-**č'er**-de-s	da-i-**č'ir**-o-s	da-e-**č'ir**-o-s
Imperfect Indicative	Conditional		Pluperfect
i-**č'er**-d-a	da-i-**č'er**-d-a		da-e-**č'ir**-a

In cases of suppletion, therefore, where the verb lexeme has three distinct stems we expect the distribution to follow Pattern B. And, indeed, for the suppletive verbs 'say' with three distinct roots ***amb-*** / *-t'q'v-* / *-tkv-*, and 'do', also with three distinct roots ***-švr-*** / *-zam-* / ***-kn-***, the distribution of the roots is in line with this morphological pattern.

Table 19: Pattern B, suppletive verbs 'say': *amb- / -t'q'v- / -tkv-*
and 'do': *švr- / -zam- / -kn-*

Series I		Series II	Series III
Present sub-series	*Future sub-series*		
Present Indicative	Future Indicative	Aorist Indicative	Perfect
amb-ob-s 'say'	i-**t'q'v**-i-s 'say'	**tkv**-a 'say'	u-**tkv**-am-s 'say'
švr-eb-a 'do'	i-**zam**-s 'do'	**kn**-a 'do'	u-**kn**-i-a 'do'
Present Subjunctive	Future Subjunctive	Aorist Subjunctive	Third Subjunctive
amb-obde-s 'say'	i-**t'q'**-ode-s 'say'	**tkv**-a-s 'say'	e-**tkv**-a-s 'say'
švr-ebode-s	i-**zam**-de-s	**kn**-a-s	e-**kn**-a-s
Imperfect Indicative	Conditional		Pluperfect
amb-obd-a 'say'	i-**t'q'**-od-a 'say'		e-**tkv**-a 'say'
švr-ebod-a 'do'	i-**zam**-d-a 'do'		e-**kn**-a 'do'

In Table 15 the presence of the thematic suffix *-eb-* for 'build' distinguishes Future Sub-series and Series III from Series II. This yields a third morphological pattern: a distinct Present Sub-series, a distinct Series II and a combined Fiture Sub-series and Series III. This third morphological pattern we will call Pattern 3. Distribution of stems according to Pattern 3 is shown in Table 20 for the verb 'build' whose preverb is *a-*.

Table 20: Morphological pattern C, 'build'

Series I		Series II	Series III
Present sub-series	*Future sub-series*		
Present Indicative a-**šen**-eb-s	Future Indicative a-a-**šen**-eb-s	Aorist Indicative a-a-**šen**-a	Perfect a-u-**šen**-eb-i-a

Since both Patterns 2 and 3 involve a three-way distinction one should expect either pattern to be followed by the suppletive items that have three distinct roots. And indeed whereas the verb 'say' above follows Pattern 2, the verb 'visit' which has three roots ***di-* / *-vl-* / *-var-*** follows Pattern 3 (Table 21).

Table 21: Pattern C, suppletive verb 'visit': *di- / -vl-*

Series I	Series II	Series III	
Present sub-series	*Future sub-series*		
Present Indicative	Future Indicative	Aorist Indicative	Perfect
di-s	i-**vl**-i-s	i-**ar**-a	u-**vl**-i-a
Present Subjunctive	Future Subjunctive	Aorist Subjunctive	Third Subjunctive
di-de-s	i-**vl**-i-de-s	i-**ar**-o-s	e-**vl**-o-s
Imperfect Indicative	Conditional		Pluperfect
di-d-a	i-**vl**-i-d-a		e-**vl**-o

So far we have met cases of suppletion where the distribution is constrained by one of three morphological patterns that exist for more regular verbs, in accordance with

Property C. A true exception, however, is the verb 'to come' which has four distinct roots *di-* / *-val-* / *-vid-* / *-sul-*. Their distribution is shown in Table 22.

Table 22: Georgian 'come' with four stems: *-di-, -val-, -vid-, -sul-*

Series I		Series II	Series III
Present sub-series	*Future sub-series*		
Present Indicative	Future Indicative	Aorist Indicative	Perfect
mo-**di**-s	mo-**val**-s	mo-**vid**-a	mo-**sul**-a
Present Subjunctive	Future Subjunctive	Aorist Subjunctive	Third Subjunctive
mo-**di**-ode-s	mo-**vid**-ode-s	mo-**vid**-e-s	mo-**sul**-iq'o-s
Imperfect Indicative	Conditional		Pluperfect
mo-**di**-od-a	mo-**vid**-od-a		mo-**sul**-iq'-o

While the division of the Present Sub-Series from the rest of the paradigm is maintained, Series III is also divided off which is not found with any of the morphological patterns. Moreover Series III and most of the Future Sub-Series are united, again not seen in the other patterns. Finally, one root is used for the Future Indicative only, again a distribution not accounted for by the posited morphological patterns.

7   Review of exceptions to the three properties.
In the previous sections we have demonstrated that in addition to high frequency there are two other properties that characterise suppletion: the suppletion involves inherent inflectional categories, and there exists a morphological pattern in a language according to which distinct stems are distributed in the paradigm. We also presented examples that do not display these properties. However, our claim was that it is the intersection of different properties that acts to preserve suppletion. In the light of this we review the exceptions to a single property and show how they follow the other two constraints.

7.1   Suppletion of an infrequent item (Property A)
As an exception to Property A we gave the Archi for 'corner of a sack', a low salience item. However, the category involved, Number for nouns, is inherent, so the example clearly adheres to Property B. We can show that it also adheres to Property C if we look at the morphological system of Archi regular nouns, together with other examples of suppletion. Archi nouns inflect for number and case. There are ten grammatical cases and a dozen locative cases. Kibrik (1977: 9) claims that a noun has four stems: singular direct, singular oblique, plural direct and plural oblique. Direct stems are based on the Absolutive, and oblique stems are based on the Ergative as shown in Table 23.

Table 23: Stem distribution of a regular Archi noun

		SINGULAR	PLURAL
a. 'ram'	ABS	baIk'	baIk'-ur
	ERG	baIk'-li	baIk'-ur-čaj
b. 'dress'	ABS	k'onc'ol	k'onc'ol-um
	ERG	k'onc'ol-a	k'onc'ol-um-čaj

In fact we will argue that there are really three distinct lexicalised stems as the fourth can be computed. The words for 'ram' and 'dress' in Table 23 use two different formatives, *-li-* for 'ram' and *-a-* for 'dress', to produce the singular oblique stem and two different formatives, *-ur-* for 'ram' and *-um-* for 'dress', to produce the direct plural stem. The same formative *-čaj-* produces the plural oblique stem for both. Altogether Archi has eight formatives available for the singular oblique stem and ten formatives for the plural direct stem, and the choice of the formative is unpredictable. But to produce the plural oblique stem, there is just one formative *-čaj-* with allomorphs *-aj-* and *-maj-* whose distribution is predictable. This means that each noun has to be learned in three forms, SgAbs, SgErg and PlAbs. Table 24 gives several suppletive examples.

Table 24. Archi suppletive nouns

		SINGULAR	*PLURAL*
a.'father'	ABS	abt:u	-
	ERG	um-mu	-
b.'man'	ABS	bošor	Lele
	ERG	bošor-mu	Lele-maj
c. 'shepherd'	ABS	ułdu	ł:ʷat
	ERG	ułdu-mu	ł:ʷa-čaj
d. 'corner of a sack'	ABS	bič'ni	boždo
	ERG	bič'ni-li	boždo-rčaj

In the first suppletive example in Table 24 'father' is a singulare tantum noun and the two distinct stems are distributed according to the distribution of singular stems of regular nouns: direct versus indirect cases. In the examples for 'man', 'shepherd' and 'corner of a sack', all of which have two suppletive stems, one stem is used for plural forms, as in the case of regular nouns. This means that the singular / plural distinction marked by the stem distribution of regular nouns is followed by suppleting nouns.

The second example of a suppletive yet infrequent item is the Hua verb 'to be heavy' in the context of Polarity. We view Polarity as an inherent verbal category: Hua verbs have a specific negative prefix *'a-* whereas nouns use a Deprivative case. Recall that the Hua verb 'be heavy' consists of two elements: *kta* 'heavy' and an auxiliary *hu* 'do' which functions as an auxiliary. There is not much inflection in Hua (it has thirty inflected verbs, all the rest being served by the auxiliaries), so there is no evidence to

establish a morphological pattern according to which the suppletion could behave. However, the verb *hu* demonstrates a kind of lexicalised patterning in the context of Polarity where /h/ alternates with /f/, since other verbs with stem initial /h/ do not show this alternation. Compare *hue* 'I do it' and *'a-fue* 'I do not do it' with *haie* 'he likes it' and negative *'a-haie* 'he does not like it' (Haiman 1980:194). Given this, we can argue that the Hua example while violating the frequency constraint adheres to both Pattern B, inherent inflection, and Property C, stem distribution that is morphologically systematic, though evidence for the latter is based on a lexically governed alternation.

7.2   Suppletion involving a non-inherent category (Property B)
The Bagwalal verb 'come'((19)-(21)) is seen as exception to Property B because the suppletion involves a Number and Gender distinction, which for verbs are contextual categories (see Table 24). However, it does adhere to Property A if we assume that salience correlates with high token frequency. Recall that cross-linguistically the verb 'come' is often suppletive: there are seven examples of suppletion for this verb in the database. The Bagwalal example also adheres to Property C in that thestems of the verb are distributed according to the existing morphological system. There are five verb conjugations in Bagwalal (Tatevosov 2001). All the verbs in the conjugation to which 'come' belongs have two stems: one for the singular and the plural non-human with formatives /a/ or /A/, and another for the human plural forms with formative /i/. Table 25 gives the distribution of stems for the verb 'bend' *aha-* / *ahi-*.

Table 25: Distribution of Bagwalal second conjugation the verb stems

SINGULAR	PLURAL

a. aram   w=aha
   person.SG   m=bend.PAST
   'a person bent'
	e. aram-di b=ahi=r
	   person-PL HUM.PL=bend.PAST=HUM.PL
	    'people bent'

b. jaš   j=aha
   girl.SG   f=bend.PAST
   'a girl bent'	

c. zin   b=aha
   cow.SG   HUM.PL=bend.PAST=HUM.PL
   'a cow bent'
	d. zin-a r=aha
	   cow-PL NONHUM.PL=bend.PAST
	   'cows bent'

In the Table we see the stem for 'bend' *aha-* alternates with *ahi-* in the context of +Human +Plural agreement. This distribution of stems corresponds to that of the suppletive example 'come', discussed in examples (19), (20), (21).

        The other example we gave of suppletion involving a non-inherent category was the Danish adjective for 'small' where two phonologically unrelated stems *lille* / *små* alternate in Number agreement, a contextual category for adjectives. Clearly the adjective belongs to the class of salient lexical items, hence it adheres to Property A. Whether or not it follows an existing morphological pattern (Property C) is less straightforward to assess since Danish adjectives distinguish Number only in some contexts: as part of an indefinite NP, or when the adjective is used predicatively. For example, compare the indefinite examples in (25) and (26) where Number on the adjective is expressed by

suffixation in the Plural, with the indefinite examples in (27) and (28), where Number is not marked on adjectives.

(25)   *en*   *glad*   *hund*
       a   happy.SG   dog.SG
       'a happy dog'

(26)   *glad-e*   *hund-e*
       happy-INDEF.PL   dog-PL
       'happy dogs'

(27)   *den*   *glad-e*   *hund*
       the.SG   happy-DEF.SG   dog.SG
       'the happy dog'

(28)   *de*   *glad-e*   *hund-e*
       the.PL   happy-DEF.PL   dog-PL
       'the happy dogs'

What is unusual about the Danish 'small' is that the distribution of the stems serves to maintain the Number distinction in both indefinite and definite contexts as shown in (10) and (11) in section 5, repeated here as (29), (30), and in (31) and (32).

(29)   *et*   *lille*   *barn*
       a   small.SG   child.SG
       'a small child'

(30)   *små*   *børn*
       small.PL   child.PL
       'small children'

(31)   *det*   *lille*   *barn*
       the   small.SG   child.SG
       'the small child'

(32)   *de*   *små*   *børn*
       the   small.PL   child.PL
       'the small children'

Looking at the nouns in the examples above we see that Danish nouns also express Number regardless of Definiteness. Given this, we must look not to the adjective but the noun paradigm to see if the suppletion example is following a pre-existing morphological pattern. And indeed for those noun lexemes where there is stem alternation, for example *barn* / *børn* 'child' as in (29) and (30) the stems are distributed according to Number, both in definite and indefinite contexts. It is this pattern that is followed by the suppletive *lille* / *små*.

7.3   Suppletion violating the pre-existing morphological pattern
The Georgian verb 'come' is seen as exception to Property C because it does not follow any of the three morphological patterns that could be said to exist in the language. However, given its lexical salience it is an item that has to be viewed as adhering to Property A. Given that the suppletion follows Tense-Aspect-Mood lines, and these are all inherent categories for verbs, it also adheres to Property B.

7.4   Summary of exceptions
We have seen evidence for three properties of suppletion. Most instances conform to all three. For example, Russian 'child' is highly frequent, the suppletion involves nominal number, an inherent category, and it adheres to an existing morphological distribution of stems that involves the sub-paradigm of number, given non-suppletive examples such as 'miracle' *čud-o* (Sg) / *čudes-a* (Pl), and 'sky' *neb-o* (Sg) / *nebes-a* (Pl). We have found some which violate one of the constraints but each one conformed to the other two. Table 26 brings together all exceptions discussed above.

Table 26: Summary of exceptions

Example	*A: High Frequency*	*B: Inherent inflection*	*C: Morphologically systematic*
Russian 'child' (normal)	✓	✓	✓
Archi 'corner of sack'	✗	✓	✓
Hua 'be heavy'	✗	✓	✓
Bagwalal 'come'	✓	✗	✓
Georgian 'come'	✓	✓	✗

8   Suppletion in copulas: the exception that proves the rule
Many languages have a suppleting copula and for this reason the copula has been a subject of interest for linguists dealing with suppletion. Most recently, Veselinova (2000) looks at the copula construction in forty-one languages belonging to twenty-two families and phyla, and extends her sample in Veselinova (2003). She shows that languages that have a regular copula tend to use it in different constructions to languages where copula is suppletive. The copula displays exceptional behaviour in relation to our constraints. Though highly frequent, therefore adhering to Property A, there are cases where the suppletion involves contextual categories of person and number, thereby flouting Property B, and at the same time following a morphological pattern that no other item in the language follows, thereby flouting Property C. For Property B English *be* is a counterexample in that it is the only verb that makes a Person distinction, contextual for verbs, by having distinct forms for 1st, 2nd, and 3rd PERSON SINGULAR: *am*, *are*, *is*. French copula *être* is also a counterexample to Property B for the same reason, and at the same time a counterexample to Property C since the distribution of the forms does not follow the pattern of regular verbs and other suppletive verbs (Table 9 in Section 6): each cell in the paradigm has a distinct form (when liaison properties are taken into account): *suis*, *es*, *est*, *sommes*, *êtes*, *sont*.

While there are several instances of copulas which are problematic for our generalization, they tend to be exceptional in other respects too. We consider frequency, overdifferentiation, fused exponence and occurrence of zeros. In terms of frequency, English *be* has a frequency which is anomalously high compared to other verbs. In the one hundred million token *British National Corpus* the verb lexeme *be* has by far the greatest frequency with 42277 occurrences of its various word-forms. This can be compared to the next most frequent verb lexeme *have* with the considerably lower frequency of 13655, and the third most frequent verb *do* with 5594 occurrences. The frequency of the copula is also anomalous when it is compared to the other major parts of speech. For nouns, *time* and *year* are the most frequent items with 1833 and 1639 occurrences respectively, and *other* and *good* are the highest ranking frequent adjectives with frequencies of 1336 and 1276 respectively. We find a similar frequency profile for the Russian copula *bit´*. In Zasorina's one million token corpus of Russian it is the most frequent verb lexeme with 13307 occurrences, where the next most frequent verb *moč´* 'be able' has about ten thousand fewer occurrences (3373), and the third most frequent is *skazat´* 'tell' with 1610 occurrences.

English *be* is also overdifferentiated in that of all the English verbs it alone distinguishes number in the past: *was* (1,3 PERSON SINGULAR), *were* (1,3 PERSON PLURAL).

Copulas often show the most severe form of suppletion, involving fused exponence. In most of the examples we have considered, items have had suppleting stems, to which irregular inflections are attached. In fused exponence, the item cannot be segmented into stem and affix (as in English *worse*, where there is no afix indicating the comparative). Similarly in the example of the French copula above, we cannot identify a stem to which affixes are attached.

A further type of irregularity involves the occurrence of zero. We find zero as a stem, in a suppletive relation to an overt stem. For example in the Basque copula *ezan* 'be' the stem *-in-/-ir-* alternates with zero in the 3rd singular past, while retaining the person, number and tense markers (Table 27).

Table 27: Zero stem suppletion in past tense of Basque copula *ezan*

PERSON	SINGULAR	PLURAL
1	n-**in**-tzen	g-**in**-en
2	h-**in**-tzen	z-**in**-en
3	z-Ø-en	d-**ir**-en

Extreme irregularity is shown by the combination of the last two problems discussed, that is a zero form in place both of stem and inflection. This type is found in the Russian copula *bit´*. It has regular forms in Past and Future but zero in the present. This is shown in Table 28.

Table 28: Zero form suppletion in Russian copula *bit′*

	Future	Past (masculine)	Present
SINGULAR			
1	bud-u	bil	Ø
2	bud-eš	bil	Ø
3	bud-et	bil	Ø
PLURAL			
1	bud-em	bil-i	Ø
2	bud-ete	bil-i	Ø
3	bud-ut	bil-i	Ø

This brief survey shows that the copula tends to be exceptional in its frequency, its marking of categories not normally marked by verbs in the language, and the nature of suppletion it displays. Given this, it is not altogether surprising that the copula is also found to be exceptional with respect to the factors presented above.

9   Conclusion.
One might expect the levelling forces of language change to reduce cases of suppletion in a given language and hence make it a relatively rare phenomenon cross-linguistically. Our lexical study of thirty languages has yielded evidence to suggest that this is not the case. If anything, finding suppletion in the lexicon is the expected situation in languages with inflectional morphology. As a step towards an explanation for why suppletion is preserved in the lexicon we have outlined a set of 'conserving' properties displayed by suppleting lexical items. This has led to a claim not about one particular property but about a group of properties that act together to lock suppletion into the lexical organization of a language. Hence the absence of one property, say high frequency, is not enough on its own to allow for reorganization of the lexeme's paradigm in favour of regularisation. Most typological claims do not rest on one factor alone but on a combination of factors, and we claim that in the case of suppletion there are (at least) three factors involved.

References
Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes.* Cambridge MA: The M.I.T. Press. (=Linguistic Inquiry Monographs 22)
Aski, Janice. 1995. "Verbal suppletion: an analysis of Italian, French and Spanish *to go*". *Linguistics* 33.403-432.
Booij, Geert. 1996. "Inherent versus contextual inflection and the split morphology hypothesis". In: Booij, Geert; and van Marle, Jaap (eds), *Yearbook of Morphology 1995* 1-15. Dordrecht: Kluwer.
Bybee, Joan. 1985. *Morphology: a Study of the Relation between Meaning and Form* (Typological Studies in Language 9). Amsterdam: John Benjamins.
Bybee, Joan. 1995. "Regular morphology and the lexicon". *Language and Cognitive Processes* 10 (5), 425-455.
Carstairs-McCarthy, Andrew. 1994. "Suppletion". In: R. E. Asher (ed.) *Encyclopedia of Language and Linguistics: 8,* 4410-4411. Oxford: Pergamon.

Clahsen, Harald; Eisenbeiss, Sonja; Hadler, Meike; and Sonnenstuhl, Ingrid. 2001. "The Mental Representation of Inflected Words: an Experimental Study of Adjectives and Verbs in German". *Language* 77 (3). 510-543.

Corbett, Greville G. 1999. "Defectiveness, syncretism, suppletion, 'deponency': four dimensions for a typology of inflectional systems". Guest lecture at The Second Mediterranean Meeting on Morphology, September 1999, Malta.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G.; and Fraser, Norman. 1993. Network Morphology: A DATR account of Russian nominal inflection. *Journal of Linguistics* 29, 113-42.

Corbett, Greville G.; Hippisley, Andrew; Brown, Dunstan & Marriott, Paul. 2001. "Frequency, regularity and the paradigm: a perspective from Russian on a complex relation". In: Bybee, Joan; and Hopper, Paul (eds), *Frequency and the Emergence of Linguistic Structure* 201-226. Amsterdam: John Benjamins.

Dressler, Wolfgang. 1985. "On the predictiveness of natural morphology". *Journal of Linguistics* 21, 321-337.

Haberland, Hartmut. 1994. "Danish." In: Kønig, Ekkehard; and van der Auwera, Johan (eds), *The Germanic Languages* 313-348. London/NewYork: Routledge.

Haiman, John 1980. *Hua: A Papuan Language of the Eastern Highlands of New Guinea* (Studies in Language Companion Series 5). Amsterdam: John Benjamins.

Hewitt B.G. 1995. *Georgian. A structural reference grammar*. Amsterdam: John Benjamins.

Hippisley 1999. Dataset of the frequencies of the noun lexeme word-forms in the Uppsala corpus. Available at: http://www.surrey.ac.uk/LIS/SMG/rusnoms.xls

Hippisley, Andrew. 2001. "Suppletion, Frequency and Lexical Storage". In: Andronis, Mary; Ball, Christopher; Elston, Heidi; and Neuvel, Sylvain (eds), *CLS 37: the Main Session. Papers from the 37th Meeting of Chicago Linguistics Society. Vol. 1* 201-14. Chicago: CLS.

Kibrik, Aleksandr E. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka: II: Taksonomičeskaja grammatika*. (Publikacii otdelenija strukturnoj i prikladnoj lingvistiki 12). Moscow: Izdatel´stvo Moskovskogo universiteta.

Leech, Geoffrey; Rayson, Paul; and Wilson, Andrew. 2001. *Word Frequencies in Written and Spoken English based on the British National Corpus*. London: Longman.

Lönngren, Lennart 1993. *Častotnyj slovar´ sovremennogo russkogo jazyka*. (=Acta Universitatis Upsaliensis, Studia Slavica Upsaliensis 33). Uppsala.

Maiden, Martin. 1992. "Irregularity as a determinant of morphological change". *Journal of Linguistics* 28, 281-312.

Maiden, Martin. 2002. "Sound change, morphemic structure and the rise of suppletion in the Romance verb". Paper presented at the *Surrey Linguistics Circle*, University of Surrey.

Mel´čuk, Igor. 1976. On suppletion. *Linguistics* 17, 45-90.

Mel´čuk, Igor. 1994. "Suppletion: toward a logical analysis of the concept". *Studies in Language* 18:2. 339-410.

Mel´čuk, Igor. 2000. "Suppletion". In: Booij, Geert; Lehmann, Christian; and Mugdan, Joachim (eds), *Morphologie / Morphology. Ein internationales Handbuch zur Flexion und Wortbildung / An International Handbook on Inflection and Word-Formation, 1. Teilband* 510-24. Berlin, New York: de Gruyter 2000.

Priestly, T. M. S. 1993. "Slovene". In: Comrie, Bernard; and Corbett, Greville G. (eds), *The Slavonic Languages* 388-451.London: Routledge.

Sparck Jones, Karen. 1999. "Information retrieval and artificial intelligence". *Artificial Intelligence* 114, 257-281.

Tatevosov S.G. 2001.Verbal morphology. In: Kibrik, A. E. (ed.), *Bagvalinskij jazyk: Grammatika: Teksty: Slovari* 67-126. Moscow: Nasledie.

Timberlake, A. 1993. Russian. In: Comrie, Bernard; and Corbett, Greville G. (eds), *The Slavonic Languages* 827-86. Routledge: London/New York.

Veselinova, Ljuba. 2000. "Suppletion in the paradigms of Copula Verbs". Fox, Barbara A.; Jurafsky, Dan; and Michaelis, Laura A. (eds), *Cognition and function in language* 65-83. Stanford: Center for the Study of Language and Information.

Veselinova, Ljuba. 2003. *Suppletion in verb paradigms: bits and pieces of a puzzle*. PhD Dissertation. Stockholm: Stockholm University.

Vincent, Nigel. 1988. "Italian". In: Harris, Martin; and Vincent, Nigel (eds), *The Romance Languages* 314-350. London: Routledge.

Zasorina, L. N. 1977. *Častotnyj slovar´ russkogo jazyka*. Moscow: Russkij jazyk.