



2001

# Suppletion, frequency and lexical storage.

Andrew R. Hippisley

University of Kentucky, [andrew.hippisley@uky.edu](mailto:andrew.hippisley@uky.edu)

**Right click to open a feedback form in a new tab to let us know how this document benefits you.**

Follow this and additional works at: [https://uknowledge.uky.edu/lin\\_facpub](https://uknowledge.uky.edu/lin_facpub)



Part of the [Linguistics Commons](#)

---

## Repository Citation

Hippisley, Andrew R., "Suppletion, frequency and lexical storage." (2001). *Linguistics Faculty Publications*. 1.  
[https://uknowledge.uky.edu/lin\\_facpub/1](https://uknowledge.uky.edu/lin_facpub/1)

This Conference Proceeding is brought to you for free and open access by the Linguistics at UKnowledge. It has been accepted for inclusion in Linguistics Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact [UKnowledge@lsv.uky.edu](mailto:UKnowledge@lsv.uky.edu).

---

**Suppletion, frequency and lexical storage.**

# Suppletion, frequency and lexical storage.<sup>1</sup>

Andrew Hippisley  
University of Surrey

## 0. Introduction

Suppletion impacts on theories of the lexicon, language change and language acquisition, and for this reason “is of considerable interest” (Mel’čuk 2000: 511). An observation often made is that there exists a relationship between suppletion and frequency (c.f. Bybee 1995, Greenberg 1996). A lexeme where there is a suppletive relationship between its word-forms usually falls within the group of higher frequency items. Because the relation is with high frequency, suppletion sheds light on the mental lexicon. High frequency words correlate with quick processing speeds, for example lexical decision times, suggesting a memory effect: the more frequently an item is accessed, the stronger the memory traces, hence the quicker the response times (see for example Clahsen, Eisenbeis, Hadler and Sonnenstuhl forthcoming). The implication is that high frequency items are stored differently. Using the results of a corpus-based analysis of Russian nouns, this paper draws together the three separate but often associated areas: suppletion, word frequency and the way items are stored in the mental lexicon. The findings show that the relationship is dependent on what is meant by frequency, which in turn suggests that suppletion can better be accounted for by combinatorial models than associative models of lexical storage.

Section 1 discusses the notion of suppletion, and using a modification of Mel’čuk’s (1994) definitional framework shows how suppletion can be viewed as a graded phenomenon covering strong suppletion, weak suppletion, and instances lying between the two extremes. Section 2 looks at what is meant by word frequency and distinguishes two kinds of frequency, *absolute* frequency and *relative* frequency. The question then is with which frequency measure is the relationship with suppletion strongest. Section 3 attempts to answer this by showing the results of an analysis of the frequencies of Russian nouns in a one million word corpus. For suppletive Russian nouns there is a stronger relationship with *absolute* frequency, with one notable exception. The implications of these results for models of the mental lexicon and lexical storage are discussed in section 4.

## 1. Suppletion

Before exploring its relationship with frequency, we need to be clear on what we take suppletion to be. Mel’čuk (1994: 343) has defined suppletion as 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.”

In Russian, stems are inflected to mark number. For example *student* : *student(y)* “student : students” where the presence of a suffix, given in parentheses, distinguishes the plural form 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 *rebënok* : *det(i)* “child : children”, where it is the stems of the word forms that 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 *rebënok* to the stem *det(i)*. On the other hand *student* : *student(y)* “student : students” must be dismissed as an example of suppletion since the relationship between the two signs is maximally regular both semantically and phonologically. We also dismiss those instances where the relationship is maximally irregular semantically but maximally regular phonologically. The word for “knee” in Russian is *kolen(o)*. It has a phonologically maximally regular relationship with the plural form *kolen(a)*, but this form does not mean “knees”, instead it means “bends in a river”.

Given the definition in (1) examples such as *child* : *children* fall outside the scope of suppletion since though the relationship between the two forms is phonologically irregular it is not “maximally irregular”. The plural stem could be analysed as *child-r(en)* where *-r-* is a stem augment and *(en)* is an irregular plural marker (c.f. *oxen*). The change in the quality of the vowel also serves to make the forms phonologically distinct. However, for Carstairs-McCarthy (1994: 4411) “suppletion is a matter of degree”, and *child* : *children* therefore counts as suppletion, albeit a weaker version of it. To incorporate such instances into Mel’čuk’s definition we can schematise it as in Figure 1.

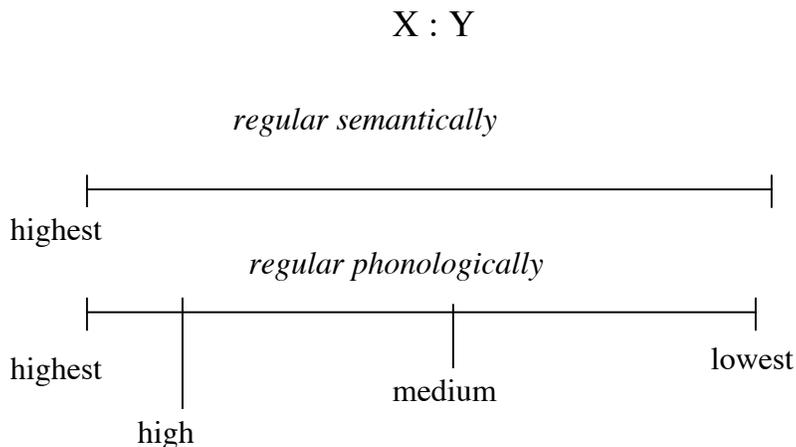


Figure 1: Suppletion as a graded phenomenon.

In Figure 1 we have two lines to denote the two dimensions of the relationship between the signs  $X : Y$ , the semantic dimension and the phonological dimension. The important point is that the lines are graded. The first line is the semantic dimension. In examples of suppletion we always use the highest setting on this line, to express semantically ‘maximally regular’. To express suppletion as a graded phenomenon, we use (at least) four settings on the line representing the phonological dimension. Examples using the different settings are given in Table 1.

Table 1: Examples illustrating the different suppletion settings  
(Note: sg = singular, pl = plural.)

<i>Setting</i>	<i>Description</i>	<i>sg:pl</i>	<i>gloss</i>
<i>Lowest</i>	full suppletion	reběnok:det(i)	‘child:children’
<i>Medium</i>	stem augments		
a.	augment in pl	neb(o):neb-es(a)	‘sky:skies’
b.	augment in sg	angličan-in:angličan(e)	‘Englishman:English’
c.	augment in sg:pl	xozjaj-in:xozjaj-ev(a)	‘landord:landlords’
<i>High</i>	stress		
a.	stem stress in sg	měst(o):mest(á)	‘place:places’
b.	stem stress in pl	okn(ó):ókna(a)	‘window:windows’
<i>Highest</i>	not suppletive	student:student(y)	‘student:students’

Prototypical suppletion, i.e. full suppletion, is expressed by the lowest setting along the ‘regular phonologically’ line (Figure 1). We have already noted the *reběnok : det(i)* example. The medium setting is used for stems that are not identical on the one hand, but not altogether dissimilar on the other. The use of stem augments in Russian yields this effect. For some lexemes stem augmentation serves to contrast singular and plural wordforms. There are three distinct augment operations. In type (a) augmentation takes place in the plural. In the example in Table 1 the formative *-es* is added to the stem *neb* in the plural to give the stem *neb-es*. In type (b) the augmented stem is the singular stem with *-in* augmenting the stem *angličan*. Finally in type (c) stem augmentation occurs in both singular and plural wordforms, but there is still a distinction between the stems since the formative used in the singular is different from the one used in the plural. For singular *-in* is attached to *xozjaj* (c.f. *angličanin*), but *-ev* is attached for the plural, *xozjaj-ev*. Note that the simple stem is *xozjaj* since it serves as the base for derivation; examples are *xozjaj-k(a)* ‘landlady’, *xozjaj-čik* ‘small landowner’, *xozjaj-sk(ij)* ‘landlord (adj)’, etc. (see Tixonov 1985). Cases where we use the

high setting are those where only stress plays a role in distinguishing singular and plural stems. There are two types. In type (a) the stem is stressed in the singular but not the plural, for example *mést(o) : mest(á)*; in type (b) the stem is stressed in the plural but not the singular, for example *okn(ó) : ókn(a)*. Finally the highest setting captures phonologically ‘maximally regular’, i.e. non-suppletive examples, and is there for completeness.

By viewing suppletion in this way as a graded phenomenon, we provide for a more fine-grained observation about the relationship between suppletion and frequency. The relationship may be stronger or weaker depending on the suppletion setting. We turn now to frequency, the second ‘partner’ in this relationship.

## 2. Frequency

By saying simply a lexeme is highly frequent, we are imprecise and in danger of being ambiguous. An item may be more frequent than another item according to *absolute* frequency, and at the same time be less frequent than the same item according to *relative* frequency. We can illustrate by comparing two noun lexemes from the Uppsala Corpus (Lönngrén 1993), reported in Hippisley (1999), given in Tables 2 and 3.

Table 2: Frequency of *Devuška* ‘girl’, Uppsala Corpus  
(Note: nom = nominative, acc = accusative, gen = genitive, dat = dative, inst = instrumental, loc = locative; sg = singular, pl = plural)

<i>singular</i>			<i>plural</i>		
Nom	devuška	57	Nom	devuški	25
Acc	devušku	18	Acc	devušek	7
Gen	devuški	22	Gen	devušek	15
Dat	devuške	8	Dat	devuškam	3
Inst	devuškoj	20	Inst	devuškam’i	6
Loc	devuške	3	Loc	devuškax	1
sg occurrences:		128	pl occurrences:		57
Total occurrences: 185					

Table 3: Frequency of *Student* ‘student’, Uppsala Corpus  
 (Note: nom = nominative, acc = accusative, gen = genitive, dat = dative,  
 inst = instrumental, loc = locative; sg = singular, pl = plural)

<i>singular</i>			<i>plural</i>		
Nom	student	12	Nom	studenti	16
Acc	studenta	2	Acc	studentov	8
Gen	studenta	0	Gen	studentov	17
Dat	studentu	1	Dat	studentam	2
Inst	studentom	7	Inst	studentam’i	8
Loc	studente	0	Loc	studentax	0
sg occurrences:		22	pl occurrences		51
Total occurrences: 73					

The frequency of the lexeme *devuška* is overwhelmingly higher than *student* if our measure is the total number of occurrences of the two lexemes. We see from the tables that the number of times *devuska* appears in the corpus is 185 compared to the 73 total occurrences of *student*. In our examples of suppletion we have contrasted the singular forms with the plural forms, hence to look for a correlation with frequency it is not the overall totals that interest us, rather the frequencies in the plural subparadigms. It is here that the ambiguity arises. When comparing *absolute* frequencies of plural occurrences, again *devuška* is more frequent than *student*: it has 57 plural occurrences compared to the 51 occurrences of *student* in the plural, bringing the lexemes much closer together, but still leaving *devuška* as the higher frequency item. However, just as interesting from the point of view of suppletion would be to look at what the 57 plural occurrences of *devuška* represent as a proportion of all its occurrences, and compare that figure with the plural proportion of *student*. To compare plural proportions amongst lexemes we are making use of a different frequency measure, *relative* frequency. This time *devuska* has an overwhelmingly lower frequency than *student*. For plurals the *relative* frequency of *devuška* is: plural occurrences/total occurrences, which is 57/185, i.e. about 31%. But for *student* the plural relative frequency is 70% (51/73). The differences in the results of the two kinds of frequency are summarised in Table 4.

Table 4: *absolute* and *relative* frequency

Plural frequencies	Devuška	Student	most frequent lexeme
<i>absolute</i> frequency	57	51	<i>devuška</i> > <i>student</i>

*relative frequency*    31%    70%    *student > devuška*

---

We have defined suppletion as a graded phenomenon and in this section have demonstrated the importance of separating out two notions of frequency. We now discuss findings of an investigation into the relationship between suppletion and frequency amongst the nouns in the 1 million word Uppsala corpus.<sup>2</sup>

### 3. Suppletion and frequency

To investigate the relationship between suppletion and frequency, we present findings of some 5500 noun lexemes and their 243000 word forms in the 1 million word Uppsala corpus. The detailed frequencies of these nouns and their wordforms are reported in Hippius (1999), an electronic dataset available on the world wide web, and discussed in Corbett et al (2001). As we have concentrated on suppletion involving the singular and plural sub-paradigms, the investigation is along the following lines. A strong relationship between suppletive lexemes and frequency will be indicated by anomalous frequency behaviour in the plural wordforms of a lexeme. In other words, with respect to the corpus at large, we are investigating whether or not the frequency of plural wordforms of what we have defined as a suppletive lexeme is anomalously high or anomalously low. Either way the anomalous frequency would indicate a relationship. We show two sets of results. In the first, the anomalous frequency we are looking for is an *absolute* frequency; in the second it is a *relative* frequency. In each case, we compare noun lexemes at the different suppletion settings discussed in Table 1 (section 1).

#### 3.1 Suppletion and *absolute* frequency

For the investigation into the *absolute* frequency, we are looking for an anomaly in the *absolute* number of occurrences of a lexeme's plural wordforms when we compare this frequency with that of the *absolute* frequency of plural wordforms of a typical lexeme in the corpus.

Table 5 shows suppletive nouns at the various suppletive settings, with a description of the setting as discussed in section 1 (see Table 1). For each setting a representative example has been chosen; the actual number of noun lexemes found in the corpus fitting the description is given in the fourth column. So for nouns fitting the Medium setting type (b) where there is an augment in the singular, ten noun lexemes were found in the corpus fitting this description. The 'M plural column' gives the median frequency of the absolute frequency of plural wordforms of all lexemes described by a given suppletion setting. For medium setting type (b), when all the plural wordforms of the ten lexemes are counted the median plural frequency is 15. Finally, and most importantly, the p-value indicates the strength of the relationship between the suppletion setting and frequency, in this case *absolute* frequency. A value < 0.05 is reasonable evidence of a relationship, and a value < 0.001 is strong evidence of a relationship.

Table 5: Russian suppletive nouns and *absolute* frequency  
(Note: M plural = ‘median plural’)

<i>Setting</i>	<i>Description</i>	singular : plural	<i>Observed types</i>	<i>M plural</i>	<i>p-value</i>
<i>Low</i>	full suppletion	rebĕnok:det(i)	3	935.5	<0.001
<i>Medium</i>	stem augments				
a.	augment in pl.	neb(o):neb-es(a)	24	10	<0.001
b.	augment in sing.	angličan-in:angličan(e)	10	15	<0.05
c.	augment in sg : pl	xozjaj-in:xozaj-ev(a)	14	14	<0.05
<i>High</i>	stress				
a.	stem stress in sg	mĕst(o):mest(á)	64	9	<0.001
b.	stem stress in pl	okn(ó):ók(na)	80	5	<0.05

### 3.1.1 Discussion

From Table 5 we see overwhelming evidence that there is relationship between a suppletive noun lexeme and *absolute* frequency, with each type displaying a significant relationship. When comparing the strength of the relationship, there is no discrimination between the different settings of suppletion, corresponding to grades between strong and weak suppletion. This is an important result for the view that suppletion is a graded phenomenon. Thus judging from the p-values, each of the settings has instances where the relationship is very strong. For the Medium setting this is the case for type (a) where the stem is augmented in the plural, e.g. *neb(o) : neb-es(a)*. And for the high setting, we find a strong relationship for type (a) where the stem is stressed in the singular wordforms, and unstressed in the plural wordforms. The low setting, full suppletion, perhaps as expected displays a strong relationship. To summarise, there is strong evidence of a relationship between suppletion and *absolute* frequency, where we view suppletion as a graded phenomenon.

### 3.2 Suppletion and *relative* frequency

To investigate an anomaly in the *relative* frequency of a lexeme, we consider the plural wordforms as a proportion of all wordforms of a lexeme and compare that figure with the plural proportion of a typical lexeme in the corpus. Examples of the suppletion settings are given again in Table 6. For the frequency column (‘M plural’) the value is a percentage to express the median proportion of plural wordforms of all lexemes described by a given setting. For example this is 20% for lexemes described by the high setting of type (a) where stress is on the stem in the singular wordforms but not on the plural wordforms. Again, evidence for a relationship is indicated by the p-values where less than 0.05 is a reasonable relationship, and less than 0.001 a strong relationship.

Table 6: Russian suppletive nouns and *relative* frequency  
(M plural = ‘median plural’)

<i>Setting</i>	<i>Description</i>	<i>singular: plural</i>	<i>M plural (%)</i>	<i>p-value</i>
<i>Low</i>	full suppletion	rebēnok:det(i)	62	<0.16
<i>Medium</i>	augments			
a.	augment in pl.	neb(o):neb-es(a)	36	<0.03
b.	augment in sing.	angličan-in:angličan(e)	82	<0.001
c.	augment in sg : pl	xozjaj-in:xozjaj-ev(a)	32	<0.4
<i>High</i>	stress			
a.	stem stress in sg	mést(o):mest(á)	20	<0.1
b.	stem stress in pl.	okn(ó):ókna)	15	<0.54

### 3.2.1 Discussion

When we compare the results from Table 6 with those with Table 5 we immediately see that the relationship between suppletion and *relative* frequency is a good deal weaker than that between suppletion and *absolute* frequency. What is interesting is that when the two tables are compared we see that the relationship is weakened for each suppletion type. The exception are lexemes described by Medium setting of type (b). Here the relationship is significantly stronger, and we discuss this finding in the next section. The general statement from these findings is that there is weaker evidence of a relationship between suppletion and *relative* frequency, including full suppletion (our ‘low setting’ type).

We can summarise our findings as follows. First, when examining the relationship between suppletion and frequency, it is important to include within the scope of suppletion examples of strong, medium and weak suppletion. This is because the relationship with frequency, at least *absolute* frequency, involves all grades of suppletion. Second, it important when claiming a relationship between frequency and suppletion that a distinction is made between *absolute* frequency and *relative* frequency. This is because the claim can only be substantiated if the relationship is with *absolute* frequency. We now look at our last area of investigation, lexical storage. The question we ask is what bearing, if any, have these findings for suppletion and frequency on the way lexemes are stored in the mental lexicon.

## 4. Suppletion, frequency and lexical storage

We have explored the relationship between suppletion and high frequency, and we now bring in our third area of investigation, lexical storage. In this section we examine what implications the findings of the suppletion and frequency study carry for two main rival models of lexical storage, the Associative Model and the Combinatorial Model.

#### 4.1 Two models of lexical storage

Clahsen et al (forthcoming) report on two main approaches to lexical storage found in the psycholinguistic literature. These are the Combinatorial Model, where (regularly) morphologically complex words are stored as stems that combine with separately stored endings in something akin to a mental paradigm. Proponents of this model include Marslen-Wilson and associates (see for example Marslen-Wilson, Tyler, Waksler and Older 1994). In the Associative Model, on the other hand, all words, both simple and complex, are stored as unanalysable wholes. Associations exist between formally similar words forms and any sense of paradigm exists merely an epiphenomenon of the associations. This idea is the inspiration behind connectionist networks such as Rumelhart and McClelland (1986). Clahsen et al use German strong adjectives to illustrate the Associative Model. Table 7 gives the strong forms of *wild* 'wild', and Figure 2 represents how these word forms are stored in an Associative Model.

Table 7: Strong declension of German *wild* 'wild'.

(Note: nom = nominative, acc = accusative, gen = genitive, dat = dative)

	<i>SINGULAR</i>			<i>PLURAL</i>
	<i>Masc</i>	<i>Fem</i>	<i>Neut</i>	<i>All genders</i>
Nom	wilder	wilde	wildes	wilde
Acc	wilden	wilde	wildes	wilde
Dat	wildem	wilder	wildem	wilden
Gen	wilden	wilder	wilden	wilder

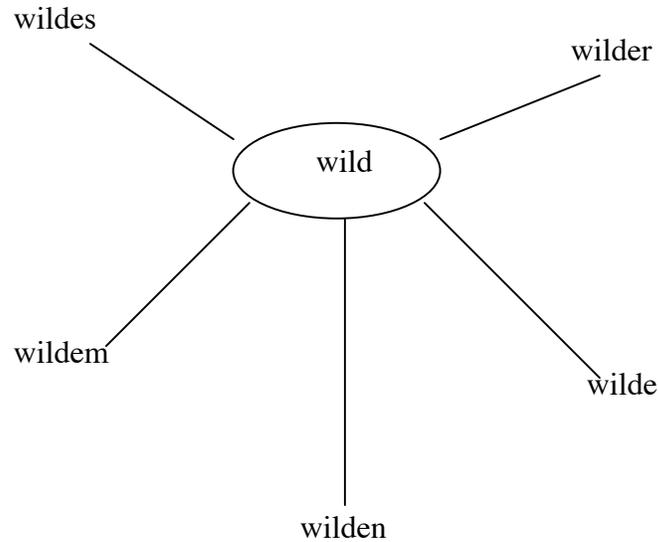


Figure 2: Wordforms of *wild* stored as wholes in the Associative Model.<sup>3</sup>

In the Associative model the basic wordform, in this case *wild*, is viewed as the *nucleus*, around which orbit the *satellite* wordforms. Any sense of ‘paradigm’ is simply the consequence of frequent associations between the nucleus and its satellites, represented by the lines.

#### 4.2 Evidence against the Associative Model

Recall from section 3 that the relationship between suppletion and frequency that has often been claimed appears to be valid when the frequency is *absolute*, but is questionable when the measure is *relative* frequency. This could be argued to constitute evidence against associative models of lexical storage. It is an interesting question how associative models which rely on formal identities to maintain connections can handle suppletion, where the formal identity is lost but the semantic identity is as strong as in the typical cases. The wordforms of a suppletive lexeme are presumably stored as in Figure 3, where x represents one form of the stem, used for singulars, and y the phonologically unrelated form of the stem, used for plurals.

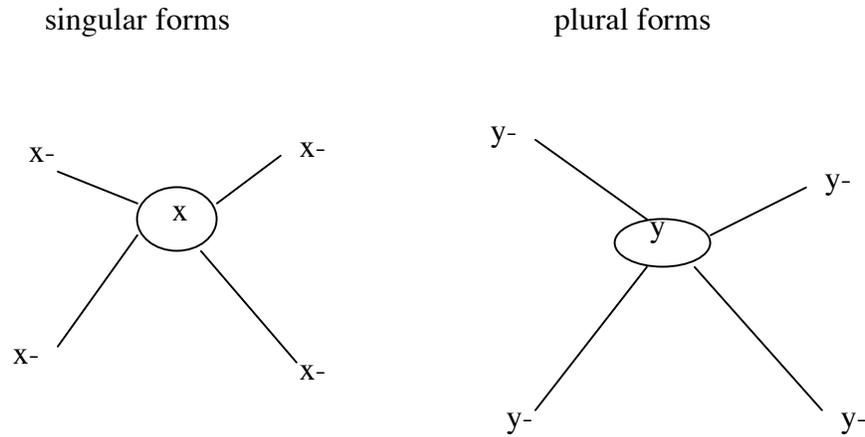


Figure 3: Suppletion in the Associative Model

In Figure 3 associations only exist between phonologically related wordforms, yielding two association groupings corresponding to the singular and plural subparadigms of a lexeme. The prediction this model has for frequency is that the frequency behaviours of the plural wordforms will differ from that of the singular wordforms. This means that singular wordforms may have either a higher frequency or a lower frequency than the plural wordforms. Whichever of these directions the frequency takes, the important point is that there will be a difference in behaviour between the two groupings of wordforms. Evidence for this from our corpus study would be that the proportion of the plural wordforms of a suppletive lexeme is significantly higher (or lower) than that of a typical lexeme in a corpus. Recall that this is expressed by any anomalies in the *relative* frequency of suppletive lexemes. But it is with *relative* frequency that we found the weakest evidence for a relationship with suppletion, including prototypical strong suppletion where there is no phonological identity at all between the wordforms. Hence the prediction associative models have for frequency effects in instances of suppletion is not borne out by the facts. Our corpus study on suppletion and frequency has therefore provided evidence in favour of combinatorial models where the paradigm of endings is primary, regardless of the shape of stems, and against associative models which rely on phonologically based associations between whole wordforms.

#### 4.3 The problem of *angl'ičan'in* : *angl'ičan(e)*

We have established that the relationship between suppletion and *relative* frequency is weaker than that between suppletion and *absolute* frequency. This is

the case for each of our identified suppletion types, with one important exception. Table 8 compares the strength of the relationship with *absolute* frequency and *relative* frequency. As can be seen from the table only for the group of nouns where the stem is augmented in the singular, while the bare stem is used in the plural, is the relationship stronger with *relative* frequency, i.e. the *angličan-in* : *angličan(e)* type.

Table 8. Comparing *absolute* frequency and *relative* frequency

<i>setting</i>	<i>description</i>	<i>singular: plural</i>	<i>strongest relationship</i>
<i>Low</i>	full suppletion	rebĕnok:det(i)	<i>absolute</i> frequency
<i>Medium</i>	augments		
a.	augment in pl.	neb(o):neb-es(a)	<i>absolute</i> frequency
b.	augment in sing.	angličan-in:angličan(e)	<i>relative</i> frequency
c.	augment in sg :	xozjaj-in:xozjaj-ev(a)	<i>absolute</i> frequency
	pl		
<i>High</i>	stress		
a.	stem stress in sg	mĕst(o):mest(á)	<i>absolute</i> frequency
b.	stem stress in pl.	okn(ó):ókn(a)	<i>absolute</i> frequency

It would appear that this suppletion type provides evidence for associative models, which would predict that the plural subparadigm behaves significantly different from the singular subparadigm, since the two subparadigms are presumably stored as two separate associations (see Figure 3). The question we now need to ask is whether this type differs in any remarkable way from all the other suppletion types. There is one characteristic that distinguishes *angličanin* : *angličan(e)* and nouns like it, namely the plural endings are irregular. The vast majority of masculine nouns whose nominative singular (citation form) is a stem lacking an ending decline like *student* in table 1. Some of these nouns take the regular endings used for neuter nouns, like *xozjaj-in* : *xozaj-ev(a)*, where the nominative and accusative (inanimate) is in *-a*. The paradigm of *angličanin* : *angličan(e)* is truly exceptional with only a handful of nouns following this pattern.

Clahsen et al. use lexical decision tasks and stem priming tasks for German verbs and adjectives to show that what may appear to be evidence for associative models is due to the one case where wordforms are stored as unanalysable wholes, namely when the ending used is irregular. One such experiment concerns German strong verbs. This is the class of verbs where the stem alternates for the present, the preterite and the participle forms, a kind of suppletion of the Medium degree. This is shown in Table 9.

Table 9. German verbs.

<i>Type</i>	<i>Present</i>	<i>Preterite (1/2 sg)</i>	<i>Participle</i>	<i>Gloss</i>
1	lauf-en	lauf-te	gelauf-t	'run'
2	sing-en	sang-Ø	gesung-en	'sing'
3	bring-en	brach-te	gebracht-t	'bring'
4	schlaf-en	schlief- Ø	geschlaf-en	'sleep'

Verbs in German can be 'irregular' on two dimensions. The first dimension is the stem, where there is alternation. The second dimension is the endings used, which are those drawn from the unproductive set of endings. Table 9 lists the various possibilities. Type 1 is the class of verbs which are regular on both dimensions. There is no stem alternation, and the endings used are the productive ones. These are *-te* for 1st and 3<sup>rd</sup> persons singular, and *-t* for the participle. The second type represents the class where there is irregularity on both dimensions. For the stem there is alternation in the present, preterite and participle forms, and the endings used are the unproductive set. This is the bare stem for the preterite 1<sup>st</sup> and 3<sup>rd</sup> singular forms, and *-en* for the participle. Types 3 and 4 are mixed in some way. For type 3 there is alternation in the stem, but the endings are regular. For type 4, on the one hand there is no stem alternation, with regard to the present and the participle; on the other hand, the participle ending is the unproductive kind.

Clahsen et al produce evidence to suggest that the only circumstance where wordforms are stored as wholes is when the ending, and not the stem, is irregular. By comparing type 2 with type 3 for preterite wordforms, where the stem alternates in both types, they show that it is the irregularity of the endings rather than the stem that produced what appears to be a lexical storage effect: the findings suggest evidence for lexical storage for type 2 which has irregular endings and not type 3, which has regular endings. Further evidence is from work done on the participle. For type 4 there is no alternation in the stem for present and participle forms, and in this regard the stem is 'regular'. But the ending used is irregular. Again there is a lexical storage effect, reported in more detail in Sonnenstuhl, Eisenbeis, and Clahsen (1999). Relating this to our findings for the *angličanin* type of suppletion which shows a strong relationship with *relative* frequency, we claim to have corroborating evidence that whole form lexical storage, even in cases of suppletion, is limited to precisely those case where the endings are irregular.

## 5. Conclusions

We started out by questioning the relationship between suppletion and frequency on the one hand, and frequency effects and lexical storage on the other. By carefully examining what can be understood by suppletion, we took the position that suppletion is a graded phenomenon that includes various degrees of 'weak' suppletion along with the prototypical strong suppletion examples. Confirmation of this is provided by the findings of an investigation into Russian nouns and

frequency, reported in an earlier paper. In the context of frequency, the various grades of suppletion display similar behaviours. We also distinguished between *absolute* frequency and *relative* frequency, and showed that evidence for a relationship between suppletion and frequency is significantly stronger when the measure is *absolute* frequency. The lack of strong evidence for a relationship with *relative* frequency was shown to disfavour associative models of lexical storage where items are stored as wholes. The one exception was a suppletion group where the evidence was stronger for *relative* frequency. This, however, turned out to be the expected situation in light of Clahsen's claim that storage of whole forms takes place when the endings themselves are irregular.

By examining the relationship between suppletion and frequency we have shed some light on how words are stored in the mental lexicon: when comparing the two major approaches to lexical storage our findings provide evidence in favour of combinatorial models where regular wordforms are analysable, and against associative models where all wordforms, irregular and regular, are stored as wholes. As has been said elsewhere, suppletion offers clues to the organisation of the mental lexicon (see Mel'čuk 2000, Carstairs-McCarthy 1994).

### Notes

<sup>1</sup> The author wishes to acknowledge the support of the Arts and Humanities Research Board.

<sup>2</sup> The results discussed in section 3 are taken directly from an earlier paper, Corbett, Hippiisley, Brown and Marriott (2001).

<sup>3</sup> Based on Clahsen et al. (forthcoming).

### References

- 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. Forthcoming. The Mental Representation of Inflected Words: an Experimental Study of Adjectives and Verbs in German. *Language*.
- Corbett, Greville; Hippiisley, Andrew; Brown, Dunstan; and Paul Marriott. 2001. Frequency, regularity and the paradigm: a perspective from Russian on a complex relation. In: Joan Bybee and Paul Hopper (eds) *Frequency and the Emergence of Linguistic Structure (TSL 45)*. Amsterdam: John Benjamins. 201-226
- Greenberg, Joseph. 1966. *Language Universals, with special reference to feature hierarchies*. The Hague: Mouton.
- Hippiisley 1999. Dataset of the frequencies of the noun lexeme wordforms in the Uppsala corpus. Available at: <http://surrey.ac.uk/LIS/SMG>.
- Lönngren, Lennart 1993. *Častotnyj slovar' sovremennogo russkogo jazyka*.

- 
- (=Acta Universitatis Upsaliensis, Studia Slavica Upsaliensis 33). Uppsala.
- Maier, Ingrid 1994. Review of Lennart Lönngren (ed.) Častotnyj slovar' sovremennogo russkogo jazyka. *Rusistika Segodnja*, 1.130-6.
- Marslen-Wilson, William; Tyler, Lorraine; Waksler, Rachelle; & Older, Lianne. 1994. Morphology and meaning in the English mental lexicon. *Psychological Review* 101.3-33.
- Mel'čuk, Igor. 1994. Suppletion: toward a logical analysis of the concept. *Studies in Language* 18:2.239-410.
- Mel'čuk, Igor. 2000. Suppletion. In: Booij, Geert; Lehmann, Christian; and Mugdan, Joachim (eds) *Morphologie*. 510-24.
- Rumelhart, David; & McClelland, James. 1986. On learning the past tense of English verbs. *Parallel Distributed Processing* 2. 216-71.
- Sonnenstuhl, Ingrid; Eisenbeis, Sonja; and Clahsen, Harald. Morphological priming and the mental lexicon: evidence from German. *Cognition* 72. 203-236.
- Tixonov, A. N. 1985. *Slovoobrazovatel'nyj slovar' russkogo jazyka*. Moskva: russkij jazyk.