Suppletion in person forms: the role of iconicity and frequency

Siewierska, A.; Bakker, D.

DOI
10.1515/9783110331127.347

Publication date
2013

Document Version
Author accepted manuscript

Published in
Languages across boundaries: studies in memory of Anna Siewierska

Citation for published version (APA):
Siewierska, A., & Bakker, D. (2013). Suppletion in person forms: the role of iconicity and frequency. In D. Bakker, & M. Haspelmath (Eds.), *Languages across boundaries: studies in memory of Anna Siewierska* (pp. 347-395). De Gruyter. https://doi.org/10.1515/9783110331127.347

Anna Siewierska and Dik Bakker[1]

Suppletion in person forms: the role of iconicity and frequency

1 Introduction

As is well known, the expression of inflectional categories in personal pronouns is often irregular. Whereas with nouns a given inflectional category may be rendered by an affix attached to a lexical stem, with pronouns often no segmentation into discernible stem and affix is possible. By way of example, compare the English *cat* vs. *cat+s* and *I* vs. *we* in regard to the expression of number. Such formal irregularity coupled with semantic regularity as in the case of *I* vs. *we* is commonly referred to as suppletion (Dressler 1985; Melčuk 1994: 358; Veselinova 2006; Corbett 2007). We too will use this term, though we are fully aware of its varying interpretations and the controversies surrounding its application with respect to oppositions in pronominal paradigms, which will be briefly discussed in Section 2.

The frequent occurrence in languages of suppletion in person paradigms is typically attributed to the high textual frequency of personal pronouns. At least since Nida (1963: 265), all forms of irregularity including suppletion have been tied to high textual frequency. The argument is not just that high textual frequency produces irregularity (though it may contribute to its emergence simply due to frequency-driven phonological erosion), but rather that it precludes or at least impedes subsequent regularization. In other words, whereas irregularity displayed by less frequent items is unstable due to the pressure of analogical levelling over generations of speakers, highly frequent items, being well entrenched and easily accessible, resist such levelling (cf. e.g. Croft 2003; Tomasello 2003; Bybee 2010). There are also scholars (e.g. Ronneberger-Sibold 1980; Werner 1987; Harnisch 1990) who adduce actual functional benefits to suppletion in high-frequency items including personal pronouns. These benefits include: the communicative advantages of short frequent forms (under the assumption that supple-

[1] This article was completed days before Anna and I went on the fateful journey from which she was not to return. It is therefore the last article of the many on which we have worked together. It goes without saying that any thought of real interest it might contain stems from her. I wish to thank the colleagues present at the Leipzig Workshop in the memory of Anna for their comments, above all Martin Haspelmath. Obviously, all remaining errors and shortcomings are mine.

tive forms are shorter), direct rather than rule-based access to frequent forms, and the perceptual advantage of maximally differentiated forms. As argued by Pike (1965), these advantages accrue to suppletion of closed as opposed to open class items since only in the former case is there no need for a productive morphological rule.

The observations made in the literature regarding suppletion in pronominal paradigms have not been confined to its frequent occurrence. It has also been noted that the distribution of suppletion differs depending on both the inflectional category and person involved. Thus suppletion is seen to be more common in the expression of number than in the expression of case, and it seems to favour the first and second person over the third (see e.g. Dressler & Barbaresi 1994; Corbett 2000: 62–66; Siewierska 2004: 48). For scholars who consider suppletion to be just a non-functional residue of diachronic change, the existence of these asymmetries in the distribution of suppletion in person pronouns is of little interest. By contrast, for those who view suppletion as potentially having a functional dimension, asymmetries in its distribution and especially the factors underlying them constitute an important research question.

While we do not exclude the possibility that the asymmetries in suppletion in personal pronouns may be due to idiosyncratic aspects of paradigmatic structure, as argued most convincingly by Maiden (2004) for the distribution of suppletion in the Romance verbal paradigms, in this article we would like to consider the case for a functional motivation of suppletion among personal pronouns. The most promising functional explanation we are aware of centres on the notion of iconicity and its various instantiations. Broadly speaking, the notion of iconicity encapsulates a correspondence between form and meaning (Peirce 1932). Within the domain of morphology, iconicity is understood as expressing the expectation that the structure of language reflects the structure of meaning in some way or other (c.f. Mayerthaler 1981; Haiman 1985a,b; Dressler 1985; Croft 2003). This basic principle has been extended in various ways to account for finer grained aspects of the semantic and formal composition of words.

An extension of special importance in the context of the present discussion of suppletion is Bybee's (1985: 24–25) Principle of Relevance, which specifies that affixes that are semantically more relevant to the meaning of the stem, should have a greater morpho-phonemic effect on the stem than the affixes which express less relevant meanings. In the case of verbal inflectional categories, for example, the Principle of Relevance predicts that fusion of the verbal stem with aspect, tense and mood affixes would be more frequent and to a higher degree than with number and person affixes. And this does indeed appear to be so (see e.g. Bybee 1985; Cinque 1999). As for the nominal inflectional categories of number and case, since number clearly affects the meaning of nominals much more strongly than does case, number being an inherent category of nominals and case possibly only a contextual one, the Principle of Relevance predicts that there should be considerably more fusion between a nominal stem and number affixes than between the stem and case affixes. When transferred to personal pronouns, the prediction thus is that suppletion in the expression of number should be more common than in the expression of case.

Bybee's Principle of Relevance is not sensitive to the lexical features of the stem; in other words it makes no predictions with respect to the lexical distribution of stem alternations with a given inflectional category. Therefore while it provides a potential explanation for the greater likelihood of suppletion in number as compared to case, it has nothing to say about the asymmetries in the amount of suppletion for first, second and third person. An explanation for such asymmetries has been sought also in the preference for an iconic relationship between meaning and form, however, in this instance with respect to the degree of transparency between the two. While transparency is normally understood as implying that transparent meanings should be encoded by transparent forms, the converse is also seen to hold, i.e. the encoding of non-transparent meaning by non-transparent form. Taking this expectation as their point of departure, Dressler & Barbaresi (1994) argue that suppletion in number should strongly favour the semantically nontransparent pairings of person and number above the semantically more transparent ones. The former, as we know, relate to the first and, to a somewhat lesser extent, second person, which in the non-singular are rarely interpreted as involving two or more speakers or hearers, respectively (Lyons 1968: 277). Rather they tend to express groups of referents which include the speaker and the hearer. The groupings associated with the non-singular first person are: speaker and addressee (1+2); speaker, addressee and other (1+2+3); or speaker and one or more others (1+3 (+3)). The groupings relevant to the second person are of the addressee and one or more others (2+3 (+3)). The third person non-singular, by contrast, is semantically much more transparent as simply more than one other is involved, just as most often is the case with non-singular NPs. Given the semantic opacity of the first and second person non-singulars and the relative transparency of the third person non-singular, Dressler & Barbaresi suggest that suppletion in number should favour the first person over the second, and both over the third.[2]

[2] Interestingly in this context, Hampe & Lehmann (this volume) observe that the frequency of the semantically 'odd' partially coreferential singular– plural pairs occurring as subject and object of the same verb, are much more frequent for first person than for second (third person pairs are not studied). This might be interpreted as support for the view that the semantic opacity with respect to singular and plural is greater for the first person than for the second.

The differences in how number is interpreted with the three persons are not paralleled by the interpretations of case. The interpretation of the accusative relative to the nominative appears to be the same for the first person as for the second as for the third. Accordingly, if Dressler & Barbaresi's explanation for the differences in the distribution of number relative to person are broadly correct, not only should there be far less suppletion in the expression of case with person than with number, as also predicted by Bybee's Principle of Relevance, but also the instances of suppletion that do occur should not exhibit the same preference for suppletion with the first person over the second over the third.

The efficacy of iconicity-based explanations for the patterns of structural encoding such as the above has been recently put into question by Haspelmath (2006, 2008), who argues that many iconicity-based explanations for asymmetrical marking patterns find a better account in terms of textual frequency.[3] As already mentioned, high textual frequency is widely recognised as the major factor underlying the frequent occurrence of suppletion in grammatical categories such as personal pronouns in general. Whether it can also be seen as underlying the discussed asymmetries in suppletion between number and case, and differences between the three persons is by no means clear. Needless to say, the possibility of a frequency-based explanation for the above is predicated on there being significant differences in frequency in the use of personal pronouns inflected for number as opposed to case, and the frequency of these inflections with first person forms as compared to second and third person ones. Interestingly enough, while the available frequency literature on the language internal use of personal pronouns reveals that there are indeed significant differences in the frequency of use of individual person pronouns, whether these differences are in line with a frequency-based account of the distribution of suppletion remains to be established.

To the best of our knowledge, neither the asymmetries in the distribution of suppletion in personal pronouns discussed above nor the viability of the functional explanations that have been invoked to account for them have ever been systematically investigated at any larger scale. The present article seeks to do so by examining in detail the distribution of number and case suppletion in free person pronouns, in particular differences between the respective persons in a cross-linguistic sample of 488 languages, and by confronting the results with both the iconicity-based and frequency-based explanations. The article is organized as follows. In section 2 we briefly review some of the controversies surrounding

[3] Actually, Haspelmath (2008) distinguishes six types of iconicity: Iconicity of quantity, complexity, cohesion, paradigmatic isomorphism, syntagmatic isomorphism, and contiguity. Only the first three are assumed to be better explained by a frequency account.

the notion of suppletion and its application to pronominal paradigms. Section 3 presents the language and areal composition of our cross-linguistic sample and of the sub-samples– languages with marked number and with marked case – that we have derived from it. Then, in section 4 we describe in detail how we have applied the typology of suppletive encoding to the person paradigms in the languages in the sample. Two methods of doing so will be described, which together should provide a robust and replicable classification of these complex data. In sections 5 and 6 we consider to what extent the asymmetries in suppletive encoding of number and case in the three persons stemming from our sample are in conformity with the iconicity-based explanations for this phenomena captured in Bybee's Relevance Principle and Dressler & Barbaresi's transparency-based (or rather: opacity-based) explanation. In section 7, we consider the data on suppletion from the perspective of the predictions following from the textual frequency of the relevant personal pronouns. As we have frequency data only for a few languages, and the interpretation of these data is far from straightforward, our comparison of the frequency-based account of suppletion in personal pronouns relative to the iconicity-based one will necessarily be more suggestive than conclusive. Finally, in section 8 we conclude the discussion with some remarks on the potential interplay between the two types of functional explanation.

2 Suppletion and person forms

Like so many other terms in linguistics, suppletion is not a homogenous notion. In traditional historical linguistics (see e.g. Rudes 1980) a distinction is made between morpho-phonologically irregular forms, which are the product of phonological change, and suppletive forms which are the result of what is sometimes referred to as incursion (see e.g. Maiden 2004: 241; Corbett 2007: 13), i.e. the invasion into a paradigm of outside forms. In other words, suppletive forms are necessarily etymologically unrelated on this older view. Nowadays, since speakers cannot be assumed to be aware of the diachronic origins of the forms they encounter, this restriction is rarely adhered to, and the term suppletion is used both for forms which are phonologically distinct by virtue of incursion and by virtue of just phonological change.[4] Needless to say, since we know next to nothing about the diachronic origins of pronouns in most languages, an investigation such as the current one is only possible if the source of the phonological irregularity of

[4] Bobaljik (2012) and presumably other adherents of Distributive Morphology are a notable exception as they treat suppletion as categorically different from irregular phonological change.

the suppletive form is not at issue. Thus, rather than viewing suppletion as combining maximal semantic regularity with extreme phonological irregularity, following Corbett (2007) and many others, we see it as part of a cline of irregularity and itself as being scalar (see further below).

2.1 Number as an inflectional category of personal pronouns

Suppletion is typically conceived of as a relation between stems within an inflectional paradigm.[5] Thus in order to be considered as suppletive the relevant forms must be involved in an inflectional alternation. There is no question as to personal pronouns being viewed as inflected for case provided the language exhibits this inflectional category. Accordingly, the English *I* vs. *me* or Polish *ja* vs. *mnie* are uncontroversially treated as suppletive. The situation with number, however, is far less clear. As discussed in the introduction, number with first and second personal pronouns is interpreted somewhat differently than with nouns. Whereas the non-singular of a noun is typically interpreted as involving more than one token of the entity denoted by that noun, and the same holds for third person pronouns, first and second person non-singulars are rarely interpreted as denoting more than one speaker or hearer, but rather receive group or associative readings. This lack of semantic transparency in their interpretation is precisely what Dressler & Barbaresi (1994) expect to be reflected iconically in their form by means of suppletion. There are scholars, however, who argue that the associative readings found with first and second person pronouns indicate that number is not an inflectional category for first and second person pronouns at all. Moreover, since, say, *I* and *we* do not express an opposition in number, the forms in question cannot be seen as suppletive. Rather, they should be viewed as distinct lexemes on a par with, say, *speaker* and *group*. A robust defence of the traditional view whereby number is an inflectional opposition within all personal pronouns is presented by Corbett (2005). We are in full agreement with his position, and mention here only two of the most important arguments for it that he cites. The first concerns formal marking, namely the fact that there are languages in which first and second person pronouns take exactly the same number affixes as third person forms and sometimes even as nouns. A case in point is that of Mizo, a Tibeto-Burman language of the Kuki-Chin group, in which the plural of all three persons consists of the singular stem with the suffix -*ni*, as shown in (1).

[5] Some scholars e.g. Mel'čuk (1994) and Markey (1985) extend the term to also include derivational and even lexical relationships (e.g. Bhat 1967).

(1) Mizo (Murthy & Subbarao 2000: 778)

	SG	PL
1	*kei*	*ke+ni*
2	*nang*	*nang+ni*
3	*ani*	*an+ni*

The second argument relates to the associative interpretations of the first and second person non-singular. Corbett points out that associative readings of non-singular forms are not in fact restricted to first and second person pronouns but rather also occur with nouns. In fact the distribution of these associative readings is governed by the position of an item on the animacy hierarchy, as first observed by Smith-Stark (1974), and subsequently documented extensively in Moravcsik (1994, 2003), and especially Corbett (2000). Thus, first and second person pronouns are not exceptional in manifesting special interpretations when inflected for number. Further, the associative interpretations found with first and second person pronouns are only one type of a range of complex readings which non-singular number may induce. In sum, there is no reason to deny the presence of a number opposition in first and second person pronouns on either morphological or semantic grounds.

2.2 Types of suppletion

The suppletion found in the marking of number and case with personal pronouns may be seen as falling into two types: total suppletion and stem suppletion. Total suppletion, which is typically considered to be the prototypical instance of suppletion, is an opposition between forms that are phonologically different from each other and are not segmentable into a stem and an affix. Total suppletion is illustrated in (2) on the basis of case marking in Polish.

(2) Polish

	NOM	ACC
1SG	*ja*	*mnie*
2SG	*ty*	*ciebie*
3SG.F	*ona*	*ją*
3SG.M	*on*	*jego*
3SG.N	*ono*	*go*

In stem suppletion, on the other hand, there is an alternation of stems which are segmentable from the affixes with which they occur. A good example of stem suppletion is given in (3) from Mangghuer, a Mongolic language spoken in China, where the addition of the plural suffix -*si* is accompanied by a change of the stem of the first person from *bi* to *da,* and of the second person from *qi* to *ta*.

(3) Mangghuer (Slater 2003: 314)

	SG	PL
1	*bi*	*da+si*
2	*qi*	*ta+si*
3	*gan*	*gan+si*

The forms involved in both total and stem suppletion may display various degrees of phonetic similarity to each other. Forms which exhibit no phonological similarity to each other at all are said to be strongly suppletive, and those which do display some phonological similarity as being weakly suppletive (see e.g. Dressler 1990: 36–37; Nübling 2000: 228; Corbett 2007).The examples of total suppletion with respect to case in Polish shown earlier in (2) and stem suppletion with respect to number in Mangghuer illustrated in (3) are clear instances of strong suppletion. It is more difficult to be confident about instances of weak suppletion without knowing well the phonological rules of a language. Since weakly suppletive stems share some common phonetic material, the possibility exists that there is a synchronic phonological rule linking the stems in question, in which case they would not qualify as suppletive. For example, in English the opposition between the strong verbs such as *think* and *thought* is typically seen as an instance of weak suppletion as there is no synchronic rule which links the two (though diachronically there is). However, rules may be devised not only for such cases but even for instances of strong suppletion, as shown by Comrie (1989). Thus, forms classified as being weakly suppletive must be viewed with some caution. A potential instance of total weak suppletion is that of the number opposition in the first person in Oromo, a widely spoken Cushitic language of Ethiopia, illustrated in (4).

(4) Harar Oromo (Owens 1985: 98)

	SG	PL
1	*na*	*nu*
2	*si*	*isi+ní*
3M	*isá*	*isáa+ni*
3F	*isíí*	

A corresponding example of weak stem suppletion is that of the case opposition in the first person in Northern Vogul, shown in (5).

(5) Northern Vogul (Riese 2001: 30)

	NOM	ACC
1SG	*am*	*an+əm*
2SG	*naŋ*	*naŋ+ən*
3SG	*taw*	*taw+e*

The distinction between strong suppletion, weak suppletion and a phonologically conditioned alternation is not always easy to draw. Strong suppletion is often associated with incursion and weak suppletion with phonologically conditioned historical changes. However, incursion can also result in weak suppletion, as in the case of Catalan, where the verb 'give' exhibits a stem alternation between *do* originating from the Latin *donare* 'donate' and *dam/dat*, originating from the Latin *dare* 'give' (Maiden 1992). And phonological change may lead to strong suppletion, as in the case of the English *am* and *is* from the Proto-Indo-European **esmi* and **esti* (Juge 1999: 186).

In section 4 we will provide a detailed discussion of how we have applied the above typology to the person forms in our sample, as well as a second, independent measure of phonological distance between two forms. To facilitate the discussion, we will assume the following terminological conventions. We will consider two forms as involving **strong suppletion** if they have no phonological material in common other than what may be taken to be coincidental. If there is some non-arbitrary phonological overlap between the relevant forms, they will be viewed as examples of **weak suppletion**. If two stems are identical, apart from possible regular phonological variation, they will be classified as **regular**. Finally, two forms may be completely **homophonous**, in which case they are classified as such.

3 The sample

In order to determine to what extent the patterns of suppletion and affixal marking involve combinations of persons consistent with the respective hypotheses under scrutiny here, personal pronouns stemming from a large sample of languages need to be considered. We collected data for a convenience sample of 488 languages of the world, which we will call S488, and which will be used as the default sample below. From this sample, using the sampling method presented in Rijkhoff & Bakker (1998) and the language classification of the Ethnologue version 15 (Ethn15; Gordon 2005) we extrapolated a subsample of 350 languages (S350) that will be employed to check the typological nature of some of our claims and observations. The languages of both samples are listed together with their genealogical affiliation in the Appendix. An idea of the makeup of the sample can, however, be deduced from the areal distribution of the languages shown in Figure 1, in which the macro-areas are essentially those used by Dryer (1989). The representation is in percentages of both samples.

In terms of the sampling technique used, the Americas and Australia are underrepresented in S488, and Eurasia, New Guinea and Southeast Asia & Oceania are overrepresented.

When we compare the distribution of the languages in both samples on the basis of the larger macro-areas proposed by Nichols (1992), we get the distribu-

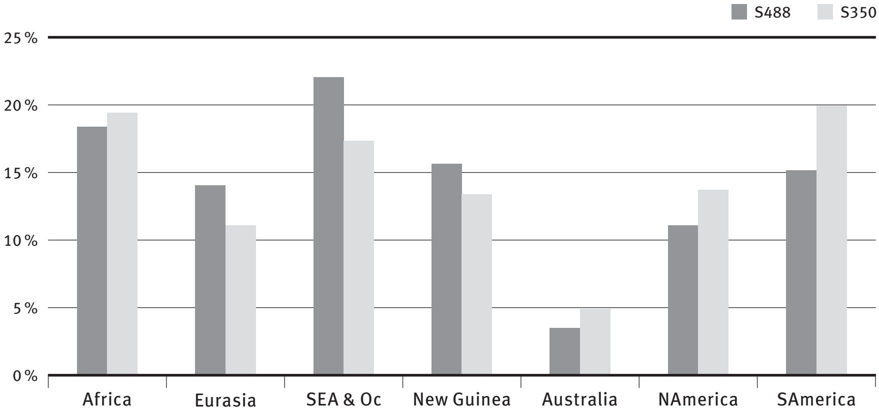

Figure 1. Distribution of the sample languages per macro-area (Dryer 1989)

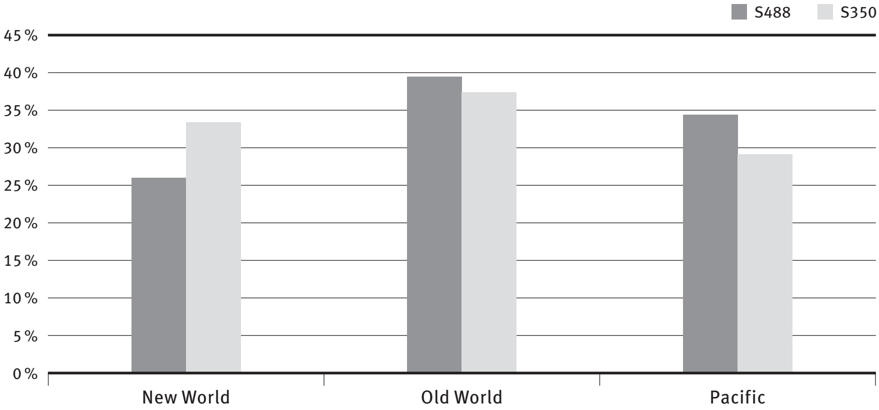

Figure 2. Distribution of the sample languages per macro-area (Nichols 1992)

tions found in Figure 2. As to be expected, we now have a clear underrepresentation of the New World, while the Old World, and more so the Pacific are overrepresented.

Of the 488 languages in the overall sample, 12 do not exhibit any number marking in their independent pronouns. Thus, for our investigation of the distribution of suppletion with respect to number the S488 sample is reduced to 476 languages, and the S350 sample to 340 languages. The final sample for case marking is considerably smaller than this. We were able to establish a differentiation between Nom/Abs and Acc/Erg in at least one of the three singular person forms for only 178 languages of S488, and for only 131 languages of S350. All the above is indicated in the language list in the Appendix.

4 The coding of suppletion

As mentioned in section 2.2. distinguishing between total and weak suppletion, on the one hand, and weak suppletion and regular phonological alternations on the other, is fraught with difficulty. A degree of arbitrariness is inevitable especially when dealing with large amounts of data, and languages with whose phonology and diachrony one is not very familiar. We have made every attempt to minimize the level of arbitrariness and maximize the degree of confidence that can be adduced to our classification of the data by employing a multi-method approach. However, before outlining the methods employed, a few words about the organization of the data are in order.

4.1 The organization of the data

For each pronominal paradigm, we entered between 2 forms (for Acoma) and 20 forms (for Ani) per language out of a theoretical total of 110, using the phonological representations of the sources consulted, typically descriptive grammars. Table 1 gives an overview of the numbers of forms selected for the respective person, number and case categories. The figures indicate the numbers of languages manifesting the corresponding forms. Alternative realizations, say of the 1SG NOM, within a language, were counted only once. When a given person-number category was further subcategorized, for example into M(asculine) vs. F(eminine) vs. N(euter), these forms were counted instead of the more general categories 1st, 2nd and 3rd. The same applies to the In(clusive)-Ex(clusive) distinction in the Pl(ural), Du(al) and Pauc(al). The empty cells in the Table 1 correspond to combinations that were not attested in any of the languages in our sample.

Table 1. Person forms in the S488 sample

		1ST	1M	1F	2ND	2M	2F	3RD	3M	3F	3N
sg	nom/abs	472	6	6	463	23	23	343	119	114	40
	acc/erg	177	1	1	175	3	3	129	48	46	17
pl	—	283	5	5	454	12	9	394	54	45	23
	In	188			2						
	Ex	184			2						
Du	—	59	7	3	107	6	6	101	12	9	2
	In	87									
	Ex	80									
Pauc	—	4	1	1	18	1	1	17	1	1	
	In	16									
	Ex	17									
Total	**4,416**	**1,567**	**20**	**16**	**1,221**	**45**	**42**	**984**	**234**	**215**	**72**

Case forms were recorded only in the singular, and only for the opposition Nom(inative)-Acc(usative) or Abs(olutive)-Erg(ative), with the absolutive corresponding to the nominative and ergative to the accusative, in accordance with markedness conventions. In the absence of a case system, the relevant form was coded as Nom. Gender distinctions other than the typical masculine vs. feminine ones were subsumed under the more common M vs. N. While for the overwhelming majority of languages we obtained complete paradigms, there were a few which had to remain incomplete.

Among the person forms, there were many with uncommon phonemes, expressed by a symbol outside the standard Ascii range as found on the regular computer keyboards.[6] The analysis program that we developed, written in a standard programming language, does not have a built-in representation mechanism for the higher (> 255) unicode characters. Therefore in order to maintain as much phonological detail as possible, we transferred the database to its hexadecimal unicode representation, substituting a phoneme by a unique number. This transformation has no influence whatsoever on the results discussed below, so we will

[6] We are aware of the fact that grammars may vary strongly with respect to the notational system used for phonological representation, especially in the case of grammars older than several decades, when typewriters and printing techniques did not always provide much opportunity for precise reproduction. Developing a unified representation system for the languages in our database was beyond the scope of the present text, and the competence of the authors. The fact that we will compare mainly forms within the same language neutralizes most of the problem, since only internal consistency counts in such a case.

continue talking in terms of phonemes. The representations in the database also contain indications of morpheme boundaries, separating stems from number and case affixes, as well as tone, coded by means of a number added to the syllable in question, as in (6) below, from Izon, a Niger-Congo language from Nigeria.

(6) Izon (Williamson 1965: 114)

SG1:	*ari5*	PL1:	*wónì4*
SG2:	*árì4*	PL2:	*ọmínì4*
SG3:	*ari5*	PL3:	*ọmínì5*

4.2 Measuring the degree of suppletion

We are not aware of any generally accepted method to measure the amount of phonological likeness between two word forms from the same language, or across languages. Therefore, in order to minimize the potential bias that relying on one method might introduce, we applied two independent, and very different strategies, one rather subjective and the other completely objective.

4.2.1 The subjective method

Applying the global suppletion scale introduced in section 2.2, we assessed the overlap for all relevant form pairs 'on face value', and from a purely synchronic angle, without taking into consideration any information about (possible) diachronic morpho-phonological developments of the languages involved. Both authors compared and coded independently of each other the degree of overlap between the first, second and third person singular and their respective plural forms, on the one hand, and the nominative and accusative singular forms for the three persons on the other hand. This was done using a system of seven codes which we applied on the basis of our linguistic intuitions. All conflicting or otherwise doubtful assignments were discussed, and a compromise was established, iteratively leading to greater consistency in the coding.

According to our coding system, a pair of forms is considered to be a case of *strong total suppletion* (TS) when the forms share no phonological material whatsoever, apart from what could be seen as a coincidence. When there is some overlap that seems to be not coincidental, then this pair is coded as showing *weak total suppletion* (TW). More often than not, this is the case when the forms share the first syllable or the first few phonemes. When one or both of the forms are morphologically marked for number and/or case we compare only the stems, in the same way that we compare full forms. When these are the same, we have *regularity* (RG). When the stems are completely different we have a case of *strong stem suppletion* (SS). And when there is some non-coincidental overlap between the two stems, we have a case of *weak stem suppletion* (SW). A special case are languages with independent person forms consisting of a person affix attached to an invariable stem as in Baruya, a Trans New Guinea language from Papua New Guinea, and exemplified in (7) below. For these forms, we ignored the invariable stem and looked only at the variable affixes, treating them like stems. So, Baruya first person is coded as SW, and second and third person as SS.

(7) Baruya (Lloyd 1989: 103)

SG1:	*ni-mino*	PL1:	*ne-mino*
SG2:	*gi-mino*	PL2:	*sari-mino*
SG3:	*ga-mino*	PL3:	*ku-mino*

In a few cases our sources did not provide enough information for us to decide whether the small differences between two stems were in fact the result of synchronically regular phonological alternations– then they would be RG – or of the insertion of epenthetic phonological material. For these we introduced the category *marginal stem suppletion* (SM). Finally, two forms may be completely homophonous (HM). Table 2 presents an example of each of the seven types of relationships between person forms that we coded.

Table 2. Coding the different kinds of suppletion

suppletion			**example**		
TYPE	**STRENGTH**	**CODE**	**LANGUAGE**	**FORM1**	**FORM2**
Total	Strong	TS	Ari	SG3.M: *nó*	PL3: *ketá*
	Weak	TW	Awa Pit	SG2: *nu*	PL2: *u*
Stem	Strong	SS	Cavinena	SG1: *ike*	PL1: *ekw-ana*
	Weak	SW	Kodava	SG2: *niini*	PL2: *nii-gal*
	Marginal	SM	Tzutujil	SG3: *jaaʔ*	PL3: *jaʔ-eeʔ*
Regular		RG	Miskito	SG3: *wĭtĭn*	PL3: *wĭtĭn-nănĭ*
Homophone		HM	Kayah Li	SG3: *ʔa*	PL3: *ʔa*

4.2.2 The Levenshtein method

Given the obvious drawbacks of the above intuitive method of classification, we introduced a second method of measuring the amount of suppletion between two person forms, this time a completely mechanical and objective, and thus fully reproducible one. This alternative method involves calculating the similarity between two forms in a language on the basis of the algorithm proposed by Levenshtein (1966). In its original, most simple form, the Levenshtein Distance (LD) between two forms is the number of steps– changes, additions, deletions – necessary to transform one row of elements (in this case: phonemes) into the other. This leads to a figure anywhere between 0 (no transformations whatsoever, complete equivalence) and n, where n is the length of the longest of the two forms (all elements transformed, maximum difference). Instead of using LD in its basic form, we adopted the normalized version as proposed by the ASJP project (cf. Brown et al. 2008; Bakker et al. 2009). This project seeks to classify the languages of the world precisely by applying the Levenshtein method to the Swadesh lists of almost 5,000 languages and dialects. The normalized version, LDN is derived from LD by dividing it by the length of the longest of the two forms, and then multiplying the result by 100, leading to a value between 0.0 (equivalence) and 100.0 (maximum distance).[7] The application of LDN to the pairs in Table 2 gives the results presented in Table 3.

Table 3. LDN values for word pairs

LANGUAGE	FORM1	FORM2	LDN	Code
Ari	SG3.M: *nó*	PL3: *ketá*	100.0	TS
Cavinena	SG1: *ike*	PL1: *ekw-ana*	66.7	SS
Awa Pit	SG2: *un*	PL2: *u*	50.0	TW
Kodava	SG2: *niini*	PL2: *nii-gal*	40.0	SW
Tzutujil	SG3: *jaaʔ*	PL3: *jaʔ-eeʔ*	25.0	SM
Miskito	SG3: *wĭtĭn*	PL3: *wĭtĭn-nănĭ*	0.0	RG
Kayah Li	SG3: *ʔa*	PL3: *ʔa*	0.0	HM

The standard LDN score, as exemplified in Table 3, is based on a segmental comparison of the two forms in question. We will label this version LDN_{phon}. For its calculation, each pair of phonemes that is compared contributes either 0 (the same phoneme) or 1 (a different phoneme) to the overall score. According to this

[7] A further operation is applied to compensate for the basic phonological overlap within a language. For details see Bakker et al. 2009.

procedure, /a/ is as different from /e/ as it is from /p/. Since this tends to overestimate the phonological difference between two forms, sometimes considerably, we added some refinement to our comparison procedure. For each of the phonemes currently found in our database we established its representation in terms of phonological features. This puts us in the position to compare pairs of feature sets rather than pairs of phonemes. Such sets will often show a certain overlap, especially in the case of two vowels or two consonants. Instead of always leading to a value 0 (equal) or 1 (different), any pair of phonemes that is compared on the basis of their respective feature sets will contribute a fraction to the total, which is rendered by the quotient (Number of features different / Total number of features compared). This expression has 0 and 1 as its limits, but it typically has a value between the two. In general, this way of measuring LDN will lead to a lower total value for the distance between two forms, which we will label LDN_{ftr}. The set of features we currently use for this operation may be found in the leftmost column in Table 4 below. It also includes some examples of the representations of four rather common phonemes. The features in Table 4 are the ones used in the P-base project (Mielke 2008). Although this set is quite basic, we believe that it provides enough dimensions for our current purpose.

Table 4. Phonological feature set and some representations

Feature	/p/	/g/	/a/	/i/
Consonantal	+	+	–	–
High	–	+	–	+
Back	–	+	+	–
Low	–	–	+	–
Round	–	–	–	–
Sonorant	–	–		
Anterior	+	–		
Coronal	–	–		
Voice	–	+		
Continuant	–	–		
Nasal	–	–		
Strident	–	–		

Using the features in Table 4 we have defined a basic list of 31 consonants and 7 vowels in terms of this feature set. In fact, this is the list in use by the ASJP project mentioned earlier. All other phonemes in the database are coded in relation to how close they are to a phoneme in the basic list. So, all phonemes in the list A = [a, à, á, â, ã, ä, å, ă, ą, ǟ, ȁ] are phonologically represented as the first element of the list, i.e. /a/. When comparing phonemes, we assign the maximum value 1.0 in case the value for their consonantal feature is different. So /p/ and /a/ will differ by 1.0. When two phonemes are in the same consonantal category, we compare the remaining set of relevant features, and determine the distance according to the formula given above. So, the distance between /p/ and /g/ is (4/11) = 0.36, and the distance between /a/ and /i/ (3/4) = 0.75. In the case of two phonemes that stem from the same set but are not identical, as for /à/ and /ã/ from set A defined above, we add one feature to the total, and stipulate 1 feature difference. Thus, all phonemes in the list A defined above will differ (1/5) = 0.2. This makes them less distant from each other than any other pair of vowels, which will score at least (1/4) = 0.25.[8]

Not surprisingly, the values for LDN_{phon} and LDN_{ftr} are rather different. For the 1,390 singular vs. plural pairs in our database, the mean value of the segment-based LDN was 60.3, while for the feature-based value it is only 38.6. Nonetheless, there turns out to be a very high correlation between the two, namely a Pearson correlation of 0.829 ($p < .01$). For the 533 pairs involving case, the correlation was even higher, at 0.900 ($p < .01$). However, it must be noted that in individual instances the values for the two measurements can be very different. Table 5 presents some striking examples.

Table 5. Differences between the LDN phoneme and feature values

Language	Form 1	Form 2	LDN_{phon}	LDN_{ftr}	LDN_{mean}
Breton	SG1: me	PL1: ni	100.0	12.8	56.4
Bukiyip	SG3.M: enan	PL3.M: omom	100.0	21.7	60.9
Amele	SG1: ija	PL1: ege	100.0	28.0	64.0
Basketo	SG1: ta	PL1: nu	100.0	38.3	69.2
Warao	SG2: ihi	PL2: yatu	100.0	49.0	74.5

The differences between the two LDN values are particularly large when the forms under comparison have the same Consonant-Vowel pattern, as in the first four examples in Table 5 above. In such cases, the feature method ‘profits’ from the C-C and V-V correspondences. In most instances the two methods lead to largely the same (relative) outcome, and it is not immediately clear which one should

[8] Other calculi have been proposed to measure the morpho-phonological distance between members of a paradigm. A very fine-grained method is introduced by Corbett et al (2001), who investigate regularity in Russian nominal paradigms. Their method contains seven levels of measurement, running from suppletion to full regularity. It assigns a greater weight to differences in stems than affixes, and counts both segmental and stress contrasts. Since it assigns weights to complete paradigms, and includes information not always available to us for all languages in our database, we think our simpler method of comparison between form pairs is to be preferred here.

be preferred. But for cases like the ones in Table 5, a more stable and realistic result can be obtained by taking the average of the two scores, i.e. the $\mathrm{LDN}_{\mathrm{mean}}$, as shown in the right-hand column in Table 5. Such a step compensates for the most extreme differences between $\mathrm{LDN}_{\mathrm{phon}}$ and $\mathrm{LDN}_{\mathrm{ftr}}$. Therefore, in the following sections we will use $\mathrm{LDN}_{\mathrm{mean}}$ as the default LDN value, while the two contributing values $\mathrm{LDN}_{\mathrm{phon}}$ and $\mathrm{LDN}_{\mathrm{ftr}}$ will be used occasionally for comparison.

As for the subjective method of classification discussed in section 4.2.1, we will regularly check the results based on the LDN values with the categories that we assigned ourselves. Some idea of the correspondences between the two systems can be gathered from a consideration of the data in Table 6, which compares the $\mathrm{LDN}_{\mathrm{mean}}$ scores for the 1,390 singular-plural pairs in our database with our subjectively-assigned categories.

Table 6. Correspondences between the two measurements: Number

Code	Pairs	$\mathrm{LDN}_{\mathrm{mean}}$	Standard Deviation	Minimum	Maximum	Range
Stem Strong	111	72.9	12.8	46.8	100.0	53.2
Total Strong	624	70.3	13.4	34.6	100.0	65.4
Stem Weak	104	40.8	13.9	18.1	84.0	65.9
Total Weak	304	37.3	12.7	9.6	78.7	69.1
Stem Marginal	28	27.3	11.0	11.5	52.5	41.0
ReGular	201	1.3	5.5	0.0	33.3	33.3
HoMophone	18	0.0	0.0	0.0	0.0	0.0
TOTAL	**1,390**	**49.4**	**28.3**	**0.0**	**100.0**	**100.0**

As Table 6 shows, there certainly does not exist a clear 1:1 relationship between the two systems in the sense of code groups corresponding to non-overlapping ranges on the $\mathrm{LDN}_{\mathrm{mean}}$ scale, with clear cut-off points. However, the mean values for LDN are very close for both types of strong suppletion (72.9 and 70.3, respectively), and for both types of weak suppletion (41.2 and 37.4), while the overall means for strong and weak suppletion (70.7 and 38.4) are a third of the scale removed from each other. The mean for the marginal stem category SM is considerably lower than the SW value, as it should be. The score for the regular cases is extremely low indeed, at 1.3, and HM, as per definition, has 0.0 for all pairs concerned. The standard deviations are relatively low, and rather constant over the categories. However, the minimum and maximum values do have a considerable overlap between the strong and weak categories. But the extreme cases– a relatively low value for 'strong' and a high value for 'weak'– are rather rare, as the standard deviations already suggest.

For the 533 pairs in our database that have a case distinction, more or less the same state of affairs holds, with the overlap between the strong and the weak categories even less than for number. We submit that these facts provide a sound basis for the use of the suppletion scale as a secondary instrument, next to the LDN_{mean}, for the analysis of the pronominal data that will be presented in section 5, to which we will turn in the next section.

At this point, however it might be interesting to compare the figures for the several types of pronominal number suppletion to those of another relatively large-scale sample on plurality in independent person forms. Daniel (2011) looked at pronominal plurality in 261 languages, of which 200 correspond to the basic sample as established by Dryer & Haspelmath eds. (2011). The comparison can only be very partial, and has to be impressionistic to some extent, since Daniel only takes into consideration first and second person forms, and in case the two plurals are derived in different ways, categorizes a language according to the way plural is coded for first person only. In his sample, 11 languages (4.2%) have no plural form, roughly the same as our figure, especially if we assume that cases of homophony are included in Daniel's percentage. Furthermore, he distinguishes six categories with respect to the way singular and plural forms differ from each other. If we generalize this to three, by skimming over the subcategories that indicate whether the plural affix is the general nominal one or a pronoun specific one, then we come to the totals in Table 7 below. Note that the percentage for Regular in the column for Daniel (2011) is a maximum, since only a subset of the languages concerned are said to have exactly the same stem for both singular and plural. No distinction is made by Daniel for the degree of suppletion, only the type.

Table 7. Comparison with data from Daniel (2011)

TYPE	Daniel (2011) (n = 250)	Our 1st Person (n = 476)
Regular	26.8 %	8.2 %
Stem Suppletion	27.6 %	17.8 %
Total Suppletion	45.6 %	74.0 %

If we assume that a certain amount of cases that are counted as Regular by Daniel would be going to (Stem) suppletion according to our standards, there is still a vast difference between the respective percentages. The relative frequencies tend to go in the same direction. However, we measure much more suppletion than Daniel does, especially of the Total kind. We assume that this is at least partially due to the way the samples are constructed.

5 Analyzing the data: Number

Our database contains a total of 4,416 paradigm slots, an average of 8.8 per language, of which we will consider only those relating to the singular vs. plural distinction. In 328 slots there were two or more alternative forms available. For these, we have decided to select the form that was presented as the first alternative in the source. In general, it is the shortest or most regular form. For each of the three persons we will use either the general Nom(inative)/Abs(olutive) form or, if there is a gender distinction, the masculine form. For the Pl(ural), we use either the single general form or, if there is a distinction in clusivity, the Ex(clusive). Of the 488 languages in our database, 448 languages (92 %) have three singular-plural pairs according to this selection criterion, one for each person. This subset, which is of particular interest here, will be called S448 below. Furthermore, 19 (4 %) languages have only two pairs, and 8 (2 %) only one pair. Finally, 12 languages (2.5 %) have no number distinction in their independent pronouns.[9] In all, we identified 1,390 Sg-Pl pairs as already presented in Table 6 above. The overall figures for the encoding of number over the three person categories are presented in Table 8. The number of different pairs (in terms of presence or absence of gender) that make up the three global person-number groups are given in the second column; we will come back to the individual pairs later.

Table 8. LDN and equivalence categories for Person-Number pairs

Person	Different Pairs SG-PL	N of pairs	Mean LDN	Homoph	Regular	Weak Suppl	Strong Suppl
1ST	5	473	56.1	.01	.07	.28	.64
2ND	5	467	50.4	.01	.14	.32	.53
3RD	4	450	41.2	.02	.23	.35	.40
TOTAL	**14**	**1,390**	**49.4**	**.01**	**.15**	**.32**	**.53**

The mean LDN value for the first person is considerably higher than for the second person, and for the latter much higher than the one for the third person. Over the whole relevant subset of languages, the three pairs of values– 1st vs 2nd; 1st vs 3rd; and 2nd vs 3rd person– correlate highly ($p < .01$). This can be summed up by saying

[9] There is one language, Ani (Khoisan; Botswana), that has both a clusivity and a gender distinction in the second person plural, and therefore would contribute four pairs, but we will select only one pair per person to facilitate the comparison between languages.

that, in terms of LDN$_{mean}$ values, pronominal number distinction is subject to the suppletion hierarchy in (8).

(8) Person1 > Person2 > Person3

This state of affairs receives support from the relative contribution per person to the four equivalence categories given in the four rightmost columns of Table 8. To determine these fractions we aggregated the figures for the three types of weak suppletion (SM/SW/TW) and those for the two types of strong suppletion (SS/TS). For all three persons the percentage of homophony is equally low, as this indeed occurs very infrequently. The amount of weak suppletion seems to be rather equal as well for the three persons. The greatest contribution to the difference between the persons stems from the low amount of regular forms for first person (7%) as opposed to the high amount for third (23%), counterbalanced by a high amount of strong suppletion for first (64%) and a relatively low amount for third person (40%). Second person takes an almost perfect intermediate position for these two categories. The bias in the distribution is significant at the .005 level, and gives independent support to relation (8). This situation is illustrated by the graphs in Figure 3 below, where we combine the scores for the four equivalence categories for each person by straight lines.

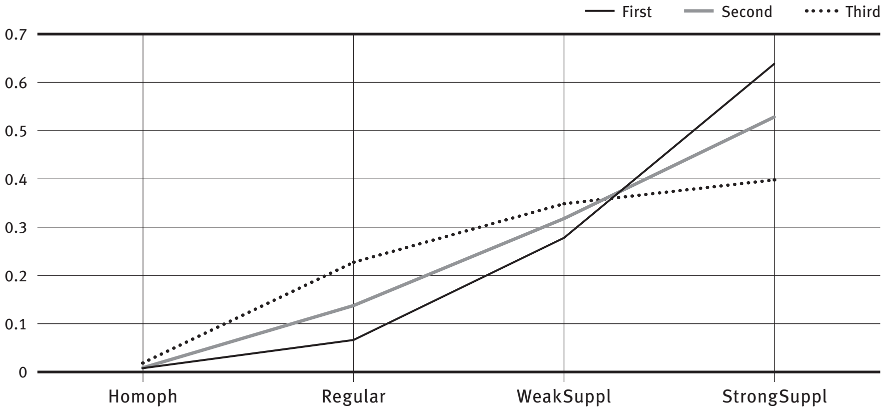

Figure 3. Graphic representation of equivalence categories for Number

Nonetheless, what we are dealing with here appears to be a tendency rather than a strict rule. The iconicity-based explanation for the distribution of asymmetries in number suppletion for the respective persons may be seen as compatible with the range of distributions captured in (9), with A corresponding to the strictest interpretation of the iconicity principle and D to the weakest.

(9) A: Person1 > Person2 > Person3 n = 98 (22%)
B: Person1 ≥ Person2 > Person3 n = 108 (24%)
C: Person1 > Person2 ≥ Person3 n = 125 (28%)
D: Person1 ≥ Person2 ≥ Person3 n = 161 (36%)

The figures on the right in (9) depict the number of languages in the S448 sample that comply with the distributions captured in A through D. We see that in 98 languages (or 22%) does the distribution of number suppletion conform to the strictest interpretation of the iconicity principle, i.e. that the first person exhibits more suppletion than the second and the second more than the third. This relatively low figure improves slightly if we allow the degree of suppletion to be equal either between the first and second person, as captured in (9.B), or between the second and third person, as captured in (9.C). Under the weakest interpretation of the iconicity principle, which allows for equality between both the first and second person and the second and third, as shown in (9.D), a little over a third of the languages in the sample (36%) are consistent with a pattern that might derive from the iconicity principle. It has to be noted that of these 161 languages, 24 have a value 0.0 for all three Sg-Pl pairs since the paradigm is regular. This leaves us with 137 (31%) 'real' cases.

The above results are somewhat surprising since the mean LDN values of the three persons given earlier looked rather suggestive with respect to there being more languages in the sample with person paradigms directly reflecting an increase in the degree of transparency in the encoding of number relative to person in accordance with the hierarchy in (8). A significantly better coverage of languages in the sample can only be achieved if we relax the interpretation of the iconicity principle even further, i.e. if we abandon the requirement that the first person should be more suppletive than the second, and juxtapose the first and second person together relative to the third, as in the two hierarchies below.

(9) E: Person1, Person2 > Person3 n = 197 (44%)
F: Person1, Person2 ≥ Person3 n = 252 (56%)

Interestingly, even now only in the case of the weaker of the two hierarchies, (9.F) which allows for any distribution of degree of suppletion in number relative to person other than that of the third person being higher than the first and second, do more than 50% of the languages in the sample comply. The only counterexamples to F are paradigms in languages such as Xokleng (Macro-Ge; Brazil) in (10), and the 'worst case' type as in Maranungku (Australian) in (11).

(10) Xokleng (Urban 1985: 167): PRS3 > PRS1 > PRS2

SG1:	*ɛ̃ŋ*	PL1:	*ãŋ*
SG2:	*a*	PL2:	*a*
SG3.M:	*tã*	PL3:	*ɔŋ*

(11) Maranungku (Tryon 1970: 16): PRS3 > PRS2 > PRS1

SG1:	*ngany*	PL1.EX:	*nga-tya*
SG2:	*nina*	PL2:	*ni-tya*
SG3:	*nankuny*	PL3:	*wi-tya*

In the sample there are 47 languages of type (10), 34 of type (11), and another 115 languages for which third person scores higher than either first or second person but lower than the other.

The above analyses nothwithstanding, the distribution of the languages in our sample over the diminishing constraint sets A to F might still be seen as providing support to an iconicity-based explanation of the formal differences between the respective Sg-Pl pairs, as long as they diverge significantly from chance distributions, and in the right direction. In order to test whether this is indeed so, we ran a large number of simulations over the same dataset, randomly drawing the LDN values for the three persons independently from the subset of 448 languages with three Sg-Pl pairs, and thus combining three pairs that originate from different languages, thus simulating a situation in which the three pairs are independent from each other. The results are given in Table 9, which also repeats the original figures from (9) for comparison. The third column gives the mean number of languages per simulation for which the relation in the first column holds.

Table 9. Comparison original sample versus random simulation

Relation	Original sample (448 languages)	Simulations (n = 30,000)	Probability Simulations > Original
A	98 (21.9%)	103.7	.709
B	108 (24.1%)	105.6	.363
C	125 (27.9%)	120.5	.304
D	161 (35.9%)	123.9	.000
E	197 (44.0%)	198.2	.525
F	252 (56.3%)	224.0	.000

We established that the means for the LDNs and the corresponding variance are virtually the same for both calculations, as of course they should be.[10] In short, the distribution of the random triplets over the six relation categories A–F does not differ much from that of the original data. In fact, in the case of two of the hierarchies –A and E– the number of languages included is even greater, and more than half of the simulations surpass the original counts in these cases. The scores for the hierarchies B and C are close, with over 30% of the simulations scoring higher. Only for the hierarchies D and F do we find considerably higher numbers for the original data. Arguably, only these two might be of further interest, with D as the most interesting candidate since it is stronger than F both statistically and in its potential implication for language. But note again that the number of languages for D would be reduced to 137 (31%) if we took out the languages with regular paradigms. Under such stricter interpretation the simulation would score such that there are more languages in the simulations than in the original data in around 5% of the cases.

For now, let us leave aside the LDN values, which have on the whole provided only partial support for the iconic basis of the distribution of suppletion over person with respect to number, and consider whether more favourable results may emerge from our subjectively determined categories. In order to determine to what extent the languages in the sample conform to the predictions of the iconicity principle in its various degrees of strength when interpreted in terms of the subjective criteria, we merged these into three individual categories: the two types of strong suppletion were collapsed into the category High, the three types of weak suppletion into Middle, and the regular cases plus the homophonies into Low. Using these three categories, the distribution corresponding to pattern A, the strongest version of the hierarchy in (8), would be H/M/L. At the other extreme, F, we would find all patterns that would have no category in any position that would be higher than its predecessor. Thus, M/M/M would fit F, but not M/L/M. Given the overlap between the respective categories this tripartite division is, of course, a less robust scale of measurement than the LDN based one. Our application of this scale to the languages in the sample with respect to each of the six interpretations A to F of the iconicity principle is shown in (12); the percentages of the LDN-based groupings are included on the right for ease of comparison.

[10] We give the figures here for interest's sake. Person1: mean 56.21– variance 639.8 (original sample) vs 56.23–636.8 (simulations); Person2: 50.02–778.9 vs 50.04–777.1; Person3: 40.99–881.6 vs 41.06–880.3.

(12)	A:	1 > 2 > 3	13 (3%)		22%
	B:	1 ≥ 2 > 3	119 (27%)		24%
	C:	1 > 2 ≥ 3	66 (15%)		28%
	D:	1 ≥ 2 ≥ 3	326 (73%)	>	36%
	E:	1, 2 > 3	123 (28%)		44%
	F:	1, 2 ≥ 3	353 (79%)	>	56%

A comparison of the two sets of data reveals that for interpretations A, C and E the subjectively based frequencies are (even) lower than the LDN-based ones. For interpretation B, which allows the first and second person to be in the same category, provided that both are in a higher category than the third person, the scores are more or less the same as for the LDN based approach. For the weaker interpretations, there is a dramatic increase in coverage when we allow the third person to be in the same group as the first and/or second (D, F). Disregarding differences between just the first and second person, as in E, has a negative effect. So, it is again interpretations F, and above all D, that stand out.

Finally, we disregard all languages that have a homophony, regularity or stem suppletion in any of the three persons, and look only at those languages that have three instances of total strong (TS) or total weak (TW) suppletion. These may be seen as the least constrained languages with respect to their person forms, since no higher order pattern is paradigmatically 'forced' upon any of them. We will call these type T languages, since all three persons have a TS or TW relation between singular and plural. This turns out to be a substantial subset, containing 240 of the 448 relevant languages (i.e. 54%).[11] However, even among these languages we do not find strong support for the iconicity hypothesis. In fact, the scores for the relations A to E turn out to be consistently lower than for the overall group, and only the weakest category F sees a higher representation for type T languages than in the random simulation.

We have seen that the distribution of suppletion in number, whether measured subjectively or objectively, does not conform to any but a rather watered down interpretation of the iconicity principle. Only in a fifth of the languages in the sample does the first person exhibit a strictly higher level of suppletion than the second, and the second person a strictly higher degree than the third. This being so, the question arises whether there is still a reason to assume that the iconicity principle plays a distinctive role in the shaping of person forms. One possi-

[11] The seven subjective categories TS, TW, SS, SW, SM, R and H as defined in section 4.2.1 could be generalized to four meta-categories T, S, R and H. Of the 64 potential patterns for the three persons that may be derived from them, only 31 occur among the 448 relevant languages in our sample, with a very skewed distribution.

bility may be that the asymmetrical distribution of suppletion in number relative to person is an areal phenomenon. Let us see whether this is indeed the case.

The distribution of the LDN values for the 448 languages with a value for all three persons relative to the three macro-areas distinguished by Nichols (1992) is presented in Table 10.

Table 10. LDN values per Macro-area (Nichols 1992)

Area	Languages	1st	2nd	3rd	Mean
NewWorld	112	51.5	34.3	28.4	38.1
OldWorld	181	58.6	52.7	43.0	51.5
Pacific	155	56.8	58.2	47.7	54.2
TOTAL	**448**	**56.2**	**50.0**	**41.0**	**49.1**

What strikes one immediately is the considerably lower mean value for suppletion in the New World. The differences are highly significant (p < .001) on a T-test for all three persons and on the overall mean in relation to the Old World, and for second and third person and the overall mean in relation to the Pacific. The Old World and Pacific differ only for the second person (p = .05). Given that the differences for the New World are mainly caused by the rather low values for second and third person, these 112 languages are potential candidates for higher scores on the relations of (9) above. This expectation is borne out by the data, be it only for relation D. The figures in Table 11 testify to this. We give both the figures for the 'inclusive' version of relation D (column two), which takes into consideration the languages with regular paradigms, and those for the exclusive version (column three).

Table 11. Areal distribution for relation D

	D (+Regular)	D (-Regular)	Total languages
New World	57 (51%)	45 (40%)	112
Old World	56 (31%)	50 (28%)	181
Pacific	48 (31%)	42 (27%)	155
TOTAL	**161**	**137**	**448**

The distribution is significant at the .005 level for the inclusive version, but only at the .05 level for the exclusive one. The inclusive version of relation D for the New World is the only relation that surpasses 50% of the relevant languages. It is the only relation that surpasses the total for random simulation in more than

99.9% of the cases (n = 30,000). A more specific test on the basis of Dryer's (1989) 7-way areal classification also shows significance at the .005 level for the inclusive version. Here, both North America (relation D applies to 57%) and South America (47%) stand out on the high side, and Africa (18%) and Australia (23%) on the low side.

Yet another check that we made is genealogical. Since the 448 languages in the database with a number form for all three persons are distributed over 109 different language families as distinguished by the Ethnologue, only a few groupings contain enough languages to enable any observations to be made. When we take into consideration only families with five or more languages in the sample, only one family stands out, namely Indo-European, with 11 out of 26 languages (42%) complying with relation A. However, this turns out to be due to genealogical overrepresentation, and a potential repetition of the same pattern. In the genealogically controlled sample of 350 languages only 4 Indo-European languages comply with relation A, which is fully in line with the overall distribution. In this subsample we found that relations A and E have scores more or less equal to the simulated set, relation B scores somewhat higher, relation C scores consistently higher with $p < .04$, and only relation D (for 38% of the languages) and relation F (55%) do better than $p < 0.005$. This may strengthen the case for D and F somewhat.

Finally, we checked the six relations A-F on the basis of the sample presented in Bybee (2005). This is a very conservative sample in the sense that great care is taken not to include languages that might derive affixes from the same etymological source, a category that is assumed to be very resistant to change. As a result, the sample contains only 26 languages, one per phylum in Voegelin & Voegelin (1977), which is taken as the guiding classification. Only six of these languages (Abkhaz, Cantonese, Guaymi, Kanuri, Karok and Tok Pisin) are also in our sample. The other 19 we replaced by the language from the same phylum in our sample that was closest to the Bybee language genealogically. Since one of these– Ojibwa, replacing Cheyenne– does not have a plural form for 2nd and 3rd person, we ended up with a sample of 25 languages. The LDN values for the three Sg-Pl pairs turn out to be even more separate from each other than for our overall sample: Person1 51.0 (overall 56.1); Person2 41.3 (50.4); and Person3 32.0 (41.2). This clearly confirms the tendency, observed for the whole database, of a greater formal distinction Sg-Pl for first than for second, and for second than for third person. Given these even greater differences, it does not come as a surprise that the percentages of languages for all relations A-F are higher for the Bybee sample than for our S448 set, and as a consequence the probabilities that a random sample scores better are considerably lower for most relations. This may be clear from the figures in Table 12.

Table 12. Comparison sample data versus random simulation

Relation	Bybee Sample (n=25)		Original data (n=448)	
	Languages	**Probability Simulations > Sample**	**Languages**	**Probability Simulations > Original**
A	7 (28%)	.165	98 (22%)	.709
B	8 (32%)	.078	108 (24%)	.363
C	10 (40%)	.162	125 (28%)	.304
D	15 (60%)	.002	161 (36%)	.000
E	12 (48%)	.082	197 (44%)	.525
F	19 (76%)	.007	252 (56%)	.000

But even if the figures for the small Bybee sample seem to be more in support of an iconicity explanation, this is still convincingly the case only for relations D and F. And even in the case of the weakest of all, F, a quarter of the languages does not follow an iconicity-based pattern. It can therefore hardly count as a universal. The strongest relations, A-C, still only apply to a minority of the languages (<< 50%).

In sum, we have seen that, across the languages in our sample, there is a strong overall tendency for a relatively high degree of suppletion between the singular and plural for first person forms, a lower degree for second person ones, and a yet lower one for third person forms. This is evident on the basis of both the 'objective' LDN approach and our 'subjective' equivalence estimate. However, these differences between the persons do not translate necessarily into implicational patterns in individual languages. The strongest version of a hierarchical relation, labelled A, applies to only 98 (22%) of the languages in our sample, and not much more in both the genealogically controlled subsample S350 (23%) and Bybee's sample (28%). Only relatively weak versions of a hierarchical relation between the three pairs, the ones we have labelled D and F, and which allow for equality between the scores for the three persons, provide stronger support for the role of iconicity in this domain. Around 36% of the languages in our S448 sample comply with relation D, 38% in our genealogically controlled S350 sample, and 60% in Bybee's 25 language sample. If we disregard the languages in our overall sample for which the paradigm is regular, we are left with around 31% for this relation. This is better than chance, and seems to suggest that iconicity could be seen as a force at work in shaping these paradigms. However, the facts and figures do not seem to be convincing enough to propose iconicity as the sole factor behind the formal differences between the three person forms with respect to number, at least not for all languages. It might however provide a partial explanation, as one of the forces at work in shaping number forms in person paradigms. Another factor might be leveling, i.e. the tendency to regularity, or frequency. Or the current situation may be determined to a high degree by phonological processes, which, over time leave more overlap between two forms in one language, where we still see correspondences between a singular and a plural form, than in another, where we observe suppletion. Such factors may play a different role in different languages and/or at different diachronic stages of the development of person paradigms, even differently for the respective person categories. Particularly illustrative of such a scenario are the examples in Nichols (this volume), who gives an in depth, historical treatment of case suppletion in person forms in several Eurasian language families. This might explain the areal and genealogical effects that we have observed.

As a final test of the potential effects of iconicity on the distribution of suppletion in number, we will compare suppletion in number to that of suppletion in case, where iconicity is not supposed to play a role of any significance.

6 Analyzing the data: Case

As already discussed in the introduction, Bybee's (1985) Principle of Relevance predicts that suppletion in case should be considerably less frequent than suppletion in number. And as we shall see below, this is indeed so. Our interest in suppletion in case, however, lies in the extent to which it displays the same person distinctions as suppletion in number. Since there have been no claims in gthe literature with respect to any iconic motivation for the existence of suppletion in case, the identification of the same person-based preferences for suppletion with case as with number would further undermine any iconicity-flavoured explanation for suppletion in number. Conversely, if suppletion in case exhibits no person distinctions, or favours the third or second person rather than the first, iconicity will emerge as a potentially credible, be it weak determinant of the distribution of suppletion in number as opposed to case. Let us see what the data hold.

As stated earlier, case is considerably less frequently expressed in personal pronouns than number. Only 178 of the languages (36.5%) in our sample have either a Nom(inative)-Acc(usative) or Abs(olutive)-Erg(ative) case opposition in their person pronouns, and even fewer languages, 165 (33.8%) have the distinction for all three persons. Taking only the singular forms into consideration, we established a total of 518 Nom-Acc or Abs-Erg pairs, selecting the forms for the masculine gender in persons manifesting such distinctions. The LDN_{mean} values per person of these forms as well as their distribution over our (generalized) equivalence categories are presented in Table 13.

Table 13. LDN and equivalence categories for Person-Case pairs

Person	N of pairs (SG)	Mean LDN	Homoph	Regular	Weak Suppl	Strong Suppl
1ST	173	39.6	.04	.36	.27	.33
2ND	171	29.9	.07	.44	.34	.15
3RD	174	26.9	.09	.45	.28	.18
TOTAL	**518**	**32.2**	**.07**	**.42**	**.29**	**.22**

A comparison of these figures with those for number in Table 8 reveals that the LDN values for case are considerably lower than those for number; the overall mean, of 32.2, is not much more than half of that for number. Thus, the Nom-Acc (and Abs-Erg) pairs are on the whole much less phonologically differentiated from each other than the corresponding Sg-Pl forms. This is confirmed by the distribution over the four equivalence categories. While for number around 85% of the pairs were at least weakly suppletive, and over 50% even strongly suppletive, for case almost half of the pairs are regular or even homophonous. The same observations hold for the individual persons. The mean LDN values are all 14 to 20 percentage points lower than for number. However, the relative order between the figures in Table 13 suggests that relation (8) from the previous section, repeated below could to some extent also be reflected in suppletion for case.

(8) Person1 > Person2 > Person3

Indeed, on a T-test, the difference for the mean LDN of case pairs between 1st person, on the one hand, and 2nd as well as 3rd on the other hand, is significant, at least at the .02 level. However, the difference between 2nd and 3rd person is not. But whatever differences there are between the persons from the LDN perspective, they almost disappear when we compare the scores for our four equivalence categories. Although 1st person scores somewhat higher in the strong suppletion column, and somewhat lower in the regular column, these differences are not significant at $p = .05$.

Given the above, there turns out to be only little reflection of relation (8) with respect to suppletion in case. This is further confirmed by the simulations for all six interpretations A–F of (9a,b), the results of which are presented in Table 14.

Table 14. Comparison of the original sample versus random simulation

Relation	Original data (165 languages)	Simulations (n = 30,000)	Probability Simulations > Original
A	22 (13%)	25.9	.814
B	28 (17%)	27.7	.459
C	38 (31%)	49.1	1.000
D	85 (52%)	62.3	.000
E	44 (27%)	43.1	.463
F	104 (63%)	94.8	.002

For four out of six relations, the scores for the data are not better – for A and C even worse – than for the random samples. The only exceptions are again relations D and F, which do considerably better, as they did for number. In this case, however a relatively high number of languages have a regular stem for all three persons.

The areal patterns follow those of number in the sense that all three areas have a pattern that globally follows (8). Again, the New World has considerably lower scores than the Old World. However, the Pacific now takes an intermediate position rather than being more or less equal to the Old World. Table 15 provides the relevant figures.

Table 15. LDN values per area (Nichols 1992)

Area	Languages	1st	2nd	3rd	Mean
NewWorld	40	27.3	21.6	16.7	21.9
OldWorld	97	48.9	34.2	32.4	38.6
Pacific	41	30.5	28.4	24.3	24.3
TOTAL	**178**	**39.6**	**29.9**	**26.9**	**32.2**

On the whole we can say that case shows the same global tendencies as number, be it in a much weaker sense than the latter, while there seems to be not much difference between 2nd and 3rd person. Furthermore, there is much more regularity for the case than we found for the number paradigm.

We may interpret the fact that there is some suppletion for case forms where we would not expect it as further weakening the position of iconicity as an explanatory factor for the distribution of suppletion in number relative to person. After all, if similar global trends are found with respect to case and to number, and no iconicity based explanation could be advanced for the former, it is not immediately clear to what extent it should be invoked for the latter. We would then have to come up with other factors that might explain the still considerable numbers of languages for which relation D holds. Alternatively, we might still maintain iconicity as an explanatory factor, under the following assumptions. Firstly, we may reassess the role of Bybee's Principle of Relevance. Counter to what we assumed earlier on, we suggest that case markers such as Nom and Acc are not merely markers of accidental contextual relations. In so-called accusative languages, Nominatives mark Subjects and Accusatives mark Objects, which are indeed syntactic categories. But in the by far most frequently occurring active sentences, the role of Subject is typically played by an Actor, and the role of Object by an Undergoer, both macroroles as defined by Van Valin & LaPolla (1997: 139f). These are semantic, not syntactic categories, and they are directly tied to the respective referents in the speech situation represented by the utterance at hand. The difference between the Actor and Undergoer functions may be so crucial pragmatically, especially when it concerns the first person, i.e. the speaker herself, that they may have given rise to specialized forms for both role types, just like non-prototypical plurality might give rise to idiosyncratic plural forms, again especially for first person. Second and third person would then be equally less sensitive for these semantic role distinctions. Mutatis mutandis for Absolutive versus Ergative. And note that in many languages, one of the pair of syntactic relations– typically 'marked' Object or Ergative– are accompanied by markers, such as adpositions. Diachronically, these may have given rise to case suffixes, which, in their turn, may have eventually lost their morphological status, leading to (weakly, then strongly) suppletive forms.

Interestingly, the suppletion phenomena are not independent. We find a rather high correspondence between number and case when we look at the rather coarse-grained grouping in High, Middle and Low suppletion, as introduced in the previous section. For all three persons, the vast majority of the languages have the same suppletion class for the number and case pairs: the χ^2 values for the distributions are significant at the .005 level. This is especially so for first person, for which also the – more precise – LDN values correlate significantly. This may be indicative of the fact that either the processes leading to suppletive forms for number and case are more or less synchronized, or that iconicity plays a more central role in certain languages or families as opposed to others.

It is obvious that this scenario would hold for only part of the languages in our sample, arguably not much more than a third of them. Therefore, there must be competition from other factors in order to arrive at the rather variegated situation testified by the figures presented above. One of the candidates for this is frequency. In the next section we will have a brief look at its potential role.

7 The frequency approach

As argued already in section 1, there is little doubt as to the role of frequency as an explanatory principle for the development and persistence of suppletion in paradigms in a more general sense. However, the issue that we would like to consider now is whether frequency underlies suppletion specifically in person pronouns and if so, to what extent it constitutes an alternative explanation to iconcity for the differences in suppletion relative to person that we have documented in section five and six.[12]

When discussing the effects of frequency on morphological form, a point of contention has been the type of frequency that needs to be considered, i.e. absolute versus relative frequency. Most studies (e.g. Schuchardt 1885; Zipf 1935; Fidelholtz 1975; Hooper 1976; Bybee & Scheibman 1999; Berkenfield 2001; and Corbett et al. 2001) consider the former to be the right choice. For example, Corbett et al. (2001) found that absolute frequencies fare much better as a predictor of irregularities in Russian nominal paradigms than relative frequencies do. They measured whether nouns for which the plural part of the paradigm showed one of eight types of irregularity, were significantly more frequent for any type in their corpus than expected. For the relative frequencies this was the case for only two irregularity types. Haspelmath (2006, 2008), on the other hand, when explaining formal differences in the coding of (in)alienability, between adjectives and their comparatives, or verbs and their derived causatives, argues that it is the relative frequencies found for the two elements of these pairs rather than their absolute token frequencies across a corpus that provide the better explanation.[13]

It seems that the two types of frequency may simply be a factor behind different kinds of phenomena. Absolute frequency is arguably behind irregularity in general, by introducing (initial) suppletion caused by the rise and maintenance of two different forms for singulars and plurals ('incursion'), and by reshaping morpho-phonologically related forms into (diachronic) suppletives over time. Relative frequency may be behind economy, with a preference for the relatively more frequent form of a Sg-Pl pair to be shorter than the less frequent one. Over time, this preference may also lead to forms becoming (diachronically) suppletive. In what follows, we will therefore consider both absolute and relative frequencies. It is important to note that there is no absolute cut off point distinguishing high

[12] Whether one sees Frequency as an explanatory factor in its own right (e.g. Haspelmath 2008) or as one instantiation of Economy (e.g. Croft 2003) is not directly relevant for the following discussion.

[13] A case for considering relative frequencies be it in connection with schemas is also presented in Hollmann & Siewierska (2007).

from low frequencies, nor whether the relative frequencies between two related forms are sufficiently different. Therefore the likelihood of suppletion developing or persisting has to be viewed in probabilistic terms.

For our comparisons we have taken into consideration only data from spoken corpora, since spoken language is the main source for the emergence and change in the shape of person forms. In performing the frequency counts we have sought to count the underlying concepts rather than the actual surface forms, in line with ideas put forward in Croft (2003: 111). Croft observes that, in discussing the singular and plural as mental constructs rather than as mere labels for linguistic forms, we should make sure that the forms considered as plural are indeed referentially plural in the specified context and not, for example, instances of pluralis majestatis or polite second person reference. Obviously, since corpus counts often span tens or even hundreds of thousands of occurrences, and the actual work is typically left to a computer program, the required type of disambiguations can only be made for fully annotated corpora. In view of the fact that only few of such corpora are currently available, we made estimates of this type of phenomenon based on 100 randomly chosen occurrences in each of the corpora that we considered. A related problem to the above, be it more tractable, is that some languages have several forms for one person-number combination while others have just one. For instance, case marking languages may have several forms for the first person singular (e.g. German: *ich* (NOM) *mir* (DAT), *mich* (ACC)) while others have just one form for all these functions (e.g. Cantonese: *ngóh*). Obviously, adding up all case forms for a person-number combination would do justice to the representation of the underlying concept (1SG), however it would overrepresent the frequency effect of the respective forms, which may in fact be very different phonologically. On the other hand, the differences between forms in a paradigm can be as subtle as the Dutch strong/weak form pairs 1SG.ACC *mij* [mɛɪ̯] vs *me* [mə], the use of which may differ only in a pragmatic sense. The most important decision that we needed to make in relation to the above was with respect to third person forms with gender distinctions. These were added up to one total for 3SG, but only provided that they were formally relatively close to each other, as for English *he* and *she*, and Spanish *él* and *ella*.[14]

[14] Experimental evidence with respect to activation, e.g. in McQueen et al. (1994) and Magnuson et al. (2007), seems to suggest that frequencies should be associated to pure forms rather than to more abstract notions such as first person. On the other hand, it is not clear to us to what extent the presence of person and number in verbal marking in the absence of free forms changes anything in the cognitive perception of the frequencies. In that respect, our figures may be both an underestimation and an overestimation. We think, however, that the corresponding frequencies are so high that the conclusions that we will base on them below will not be affected by these choices in a fundamental sense.

Table 16 presents the frequency data for 12 languages, with the sources from which they were derived. All these sources are spoken corpora available on line. Full references may be found at the end of the article.[15] We are aware of the fact that most of the languages in question are genealogically related, and do not represent independent cases in the typological sense. However, building a collection of relatively large spoken corpora for a sample of languages equivalent to the one we used for our typological exercise above, though a highly desirable goal, seems to be illusionary at the moment.

Since the corpora are very different in overall size, we give the frequencies in terms of occurrences per 1000 tokens, which is roughly equivalent to 5 minutes of spoken discourse. The PD column in the table indicates whether a language is fully pro-drop (Y), partially (P), or not (N).

Table 16. Corpus totals of pronouns per 1000 tokens: Number

Language	PD	Source	Tokens $(x10^6)$	1SG	1PL	2SG	2PL	3SG	3PL	TOTAL
English	N	BNC	4.2	40.2	8.0	20.4	11.7	19.5	10.5	110.3
Estonian	Y	MDC	.1	12.0	2.3	3.4	3.1	6.5	2.2	29.5
French	N	Beech	.1	17.4	3.4	9.1	5.5	19.2	6.3	60.9
German	N	DGD	.2	18.5	6.7	1.2	0.6	11.6	5.7	44.3
Hebrew	P	HSC	.03	29.1	3.1	16.5	1.7	21.7	6.0	78.2
Italian	Y	BADIP	.5	7.7	2.8	2.1	1.1	3.2	1.2	18.0
Polish	Y	Pelcra	.6	9.4	0.7	1.9	0.3	5.8	1.5	19.6
Portuguese	Y	Davies	1.1	18.3	3.4	0.6	0.1	4.0	0.8	27.1
Russian	P	RNC	9.6	9.7	2.8	3.2	3.5	5.2	6.0	30.3
Scots	N	SCOTS	1.1	33.7	8.9	15.7	8.8	9.8	7.9	84.8
Spanish	Y	Davies	5.1	6.0	1.0	1.1	0.1	1.9	1.0	11.1
Swedish	N	GSCL	1.4	21.3	14.5	12.1	2.0	5.7	0.7	56.2

Looking at these figures, what strikes one first are the considerable differences between the languages in the *absolute* frequencies of the respective pronouns. As to be expected, the main explanation for this turns out to be the pro-drop factor. We measured a mean of 71.1 pronouns per 1000 tokens for the non-pro-drop languages, 54.3 for the partial ones, and 21.1 for the fully pro-drop languages in the

[15] We are extremely grateful to the many colleagues out there on the web who were very helpful with providing access to these corpora, and with searching them. Their amicable attitude turned out to be exemplary.

corpora, i.e. around 7%, 5% and 2% of the running texts, respectively.[16] The greatest overall difference is between English with 11%, and Spanish with just over 1%. But even in the latter language, all three singular forms were found among the 100 most frequent types, all with a frequency higher than 1 per 1000 tokens. It seems to be safe to consider at least the singular forms as 'frequent' in the absolute sense for all languages concerned, since they occur at least once every few minutes of spoken discourse.

Of the singular forms, 1SG is by far the most frequent for all but one language, followed by 3SG in most of the cases, and then 2SG. For French, the order is 3SG > 1SG > 2SG. For the plural forms, we generally find the same order of the persons as for the singular. The only language that diverges more or less clearly from this pattern is Russian, with high 3PL and low 2SG values. Arguably, this is caused by the very frequent use of the third person plural form as an impersonal, possibly at the cost of the 2SG in that same function (cf. Siewierska & Papastathi 2011). We have made no attempt here to distinguish between the personal and impersonal usage of pronouns, a phenomenon that occurs in most languages, however not necessarily to the same extent for the same person-number combinations.

If we could accept that there might be minor divergences for individual languages, possibly based on differences in the way person forms are used for other functions, such as impersonality, but that do not fundamentally affect global tendencies, then the following frequency hierarchy would hold:

(13) a. 1SG > 3SG > 2SG > 1PL > 3PL > 2PL

This implies (13b) and (13c):

(13) b. SG > PL
c. 1 > 3 > 2

If frequency were to be the major factor determining suppletion, then the amount of suppletion among the six forms across the languages in our sample of sections 5 and 6 above should echo the hierarchies in (13). In order to establish this, we turned back to our database of section 5, and calculated the LDN values for each

[16] The only outlier seems to be Hebrew, a language with (partial) pro drop, and which scores higher than non-pro drop French and German. But with only 26,500 tokens, the Hebrew corpus is by far the smallest, which makes statistical observations relatively unreliable. Furthermore, in general, differences may be caused by the nature of the corpora. Although all are spoken, there may be considerable differences in the use of pronouns between, e.g. informal telephone conversations and television interviews. We have not controlled for type of corpus. And there may be other factors that might create considerable differences, e.g. cultural ones.

of the six forms with respect to all the other forms in the corresponding paradigms. This provides us with a relative measure for how each form stands out within its paradigm. We found the following (the mean LDN values for the 476 languages in the sample are in brackets):

(14) 1SG (57.4) > 1PL (56.6) > 2SG (56.4) > 3SG (55.8) > 3PL (55.4) > 2PL (55.1)

The differences between the mean LDN values are rather small. Still, the ranking in (14) falls largely in step with the relative amounts of suppletion we found for the three person-number pairs in section 5, especially when we add up the values for the three persons. This would give us the relation 1 > 2 > 3 rather than the frequency based 1 > 3 > 2 order in (13c). Our conclusion must be that frequency can not be the sole, or even the most important explanatory factor for the differences in the amount of suppletion that we find for the six pronouns.

Let us now turn to the person-number pairs, and look at their *relative* frequencies. We first have to establish what we would predict in terms of suppletion as a result of differences in relative frequency, ignoring for the occasion the potential role of absolute frequency. Our assumption is that one would expect more regularity when the relative frequencies are very different, with the most frequent form of the two as the least marked, or ‘zero’ form, and the least frequent form the one with the additional morphology. Conversely, if suppletion would be motivated by relative frequency at all, one would expect it in case both forms occur more or less equally frequently. When we look at the actual relative frequencies for the three person-number pairs we find the proportions given in the second column of Table 17 below. The Relative Frequency Quotient (RFQ) in column two in the table is calculated as follows.

(15) $\text{RFQ} = ((f_H / (f_H + f_L)) - 0.5) * 2 \; [\, f_H > 0 \,]$

In (15), f_H is the frequency found for the most often occurring element of the pair in some corpus, and f_L the frequency of the least often occurring element. The minimum RFQ value of 0.0 indicates equal frequencies. The maximum, 1.0, is reached in case one of the two frequencies is 0. The differences between the 12 languages for which we have corpus counts are considerable, especially for second person, as the minimum and maximum values in columns three and four show. For a somewhat more balanced impression of the mean proportions, we have excluded the outliers, i.e. scores more than 1 S(tandard) D(eviation) from the mean.

Table 17. Relative frequency quotients Singular vs. Plural

Person	Mean RFQ	Minimum	Maximum
1	.57	.19 (Swedish)	.86 (Polish)
2	.69	.04 (Estonian)	.96 (Portuguese)
3	.47	.07 (Russian)	.78 (Swedish)

The figures in Table 17 would predict that we should find most suppletion for third person, less for first, and least for second, i.e. 3 > 1 > 2. If anything at all, our measurements of the suppletion levels in section five showed a tendency towards 1 > 2 > 3. The same rather negative result is found when we correlate the relative frequencies of the three person-number pairs with the corresponding LDN values for each individual language. If our hypothesis about the relation between level of suppletion and relative frequency held, then there should be a (significant) negative correlation between LDN and RFQ. Nothing of the kind was found in our data.[17] Thus, also relative frequencies do not seem to be a very convincing factor for determining the amount of suppletion among the person-number pairs of independent pronouns.

Finally, we searched for frequency effects with respect to the case forms. The corpora provided us with the figures in Table 18. As in Table 15 above, the unit is occurrences per 1000 tokens.

Table 18. Corpus totals of pronouns per 1000 tokens: Case

Language	PD	1SG	1SG.ACC	2SG	2SG.ACC	3SG	3SG.ACC
English	N	40.2	3.9	20.4	2.1	19.5	4.7
Estonian	Y	12.0	1.0	3.4	0.6	6.5	0.3
French	N	17.4	3.2	9.1	0.7	19.2	0.8
German	N	18.5	1.3	1.2	0.1	11.6	0.5
Hebrew	P	29.1	0.8	16.5	0.9	21.7	3.7
Italian	Y	7.7	4.4	2.1	2.2	3.2	7.0
Polish	Y	9.4	2.1	1.9	0.8	5.8	2.3
Portuguese	Y	18.3	4.2	0.6	0.3	4.0	9.6
Russian	P	9.7	2.0	3.2	0.8	5.2	2.8
Scots	N	33.7	3.0	15.7	1.6	9.8	2.6
Spanish	Y	6.0	5.2	1.1	1.8	1.9	5.8
Swedish	N	21.3	1.4	12.1	1.2	5.7	0.5

[17] The correlations that we found were never significant, and even slightly positive in two out of the three cases: first person (Pearson/Kendall): .000 / .078; second person: .196 / .023; third person –.080 /–.156.

We detected the following tendencies:

(16) a. 1NOM > 3NOM > 2NOM > 3ACC > 1ACC > 2ACC
b. NOM > ACC
c. 1, 3 > 2

Exceptions to these implications are the predominant ACC > NOM order for the third person of the three Romance languages, and for the second person in the case of two of these, Spanish and Italian. Although a higher relative frequency for accusatives might be expected for pro-drop languages in general, and indeed is found also for the two Slavic languages, an ACC > NOM order could not be attested anywhere else in the corpora. The frequency orders of (16) would suggest suppletion values vis à vis the total paradigms for the forms in the same order, i.e. most suppletion for 1NOM and least for 2ACC, and more for NOM than for ACC. The following, however, was found for the case marking languages (the LDN values are again in brackets):

(17) 1ACC (59.9) > 3ACC (58.1) > 2ACC (57.7) > 1NOM (57.5) > 2NOM (56.5) > 3NOM (55.8)

This is clearly not a confirmation of the frequency-based order in (16a), and it definitely does not give any support to (16b). Note that the latter can not be caused by the influence of regular case markers, since only stems are compared in such cases.[18]

Following the same procedure as for the number pairs, we also established the relative frequencies for the three person-case pairs. We found the following RFQ values: 1NOM.ACC = .76 > 2NOM.ACC = .66 > 3NOM.ACC = .58. This would predict precisely a reverse 3 > 2 > 1 order in suppletion. As shown in section 6, this is certainly not the order we found for the overall LDN scores. When we look at the values for the 12 languages for which we have corpus data, we do find negative correlations for all three persons, but all of these have probability values way above the 5% level. So, also suppletion in the pronominal case forms defies a convincing frequency-based explanation.

We have to conclude from this that frequency can not be the single explanation behind the distribution of suppletion that we found for the person forms in

[18] Another, quite plausible explanation, pointed out by Martin Haspelmath (p.c.), may be the generally accepted fact that first and second person forms tend to be much older in a language than third person forms, which are more often 'recycled', typically by reinterpretation of demonstratives. As a result, they may show less suppletion than second person forms, despite the fact that they are more frequent in discourse.

our sample, not even for the small subset of languages for which we have corpus data. It may, however, be a force in competition with others, arguably iconicity. We have argued that, for number, the relative frequencies would predict third person forms to have most suppletion, and second person least, i.e. more or less in the opposite direction that iconicity was assumed to work. So, we may have more suppletion for third person than we would have if iconicity were the only factor, and less for second and first person.

This would then explain the fact that, although there is a weak global tendency towards 1 > 2 > 3, there is a lot of variation among the languages in our sample. There are even a few languages that go completely counter to that tendency. With iconicity as the most prominent force overall and frequency as a secondary one, we may have a better explanation for the distribution of suppletion that we found in our sample. For the further fine-tuning of our understanding of this phenomenon we may have to appeal to factors such as pro-drop, the impersonal use of certain person forms, the existence and frequency of use of a passive, medium or inverse, the morphological type, and yet others, such as the push to regularity. Since, with suppletion, we are mainly looking at the results of processes in the (remote) past, diachronic rather than synchronic information about the characteristics of the languages concerned is called for.

8 Concluding remarks

Our investigation of the distribution of number and case suppletion in person paradigms has shown that while there is indeed a global preference for suppletion to favor the first person over the second over the third, this is certainly not a universal in the strict sense of a typological hierarchy, or a Greenbergian kind of implication. At best, what we observe is a tendency, with many (apparent) counterexamples. But for a minority of the languages the first person does indeed exhibit more suppletion than the second and the second more than the third, 22% of the languages under a strict interpretation of such a distribution, and 36% under a more liberal view, which cannot simply be ignored, or attributed to chance. Our findings are thus reminiscent of those of Bybee (2005), who documents that there is indeed a cross-linguistic tendency in languages to use a restricted set of phonemes, especially unmarked ones, in inflectional affixes as opposed to stems, but only a weak one. Just as in our case of suppletion, the majority of the languages in her sample did not display the relevant patterning despite the unequivocal existence of such a patterning in some languages. In short, what is purported to be a universal may rather be interpreted as a tendency, apparent in some languages while obscured by competing factors in other languages. Significantly, only investigations as detailed as Bybee's (2005) and Nichols (this volume) can hope to determine what is actually going on. Ideally, we would also want to have access to a representative corpus of spoken discourse for each of the languages in our database, and for which we collected paradigmatic data. And ideally, our knowledge about the diachronic stages of the languages concerned, and above all of the history of the person forms, would be far superior to what we know today.

Our findings are also reminiscent of Bybee's with respect to the role that functional factors are likely to play in determining the existing distributions of the relevant phenomena. Bybee concludes that the distribution of phonemes in inflectional affixes vs. stems in her data is not amenable to any single explanation but rather must be attributed to multiple diachronic trends such as phonological reduction in grammaticalization, and the re-use of old affixes in creating new ones. We, too have argued that neither of the two functional factors that have been invoked in the context of suppletion with respect to personal pronouns, iconicity and frequency, suffice to account for the distribution of suppletion in person paradigms that we have documented.

The primary functional factor that we considered was iconicity, more specifically the hypothesis that the greater semantic opacity of the first person nonsingular as compared to the second, and the second as compared to the third should be accompanied by corresponding greater morphological opacity, and thus a greater likelihood of suppletion of the first person relative to the second, and of the second relative to the third. Although the global, be it weak, trend that we found is in line with the predictions of iconicity, the same order of relative suppletion, be it weaker version, was observed in the nominative-accusative pairs of the singular forms. In this case, no comparable iconicity-based explanation is generally assumed to be applicable. We, however speculated that the role of Relevance may be invoked here as well, which would give some extra support to the iconicity hypothesis.

The other functional factor potentially underlying suppletion that we examined was frequency. The hypothesis that the most frequent forms would be also the ones most likely to evince suppletion turned out to be even weaker than the iconicity hypothesis. Although absolute frequencies do give support to the fact that most suppletion is found in the first person singular, they would further predict that third person would be more suppletive than the second person, both in the singular and the plural, which was not what we had found on the basis of our two approaches to comparing the forms. And relative frequencies for the three person-number combinations would predict that third rather than first person would be the most suppletive pair, with second person coming last, which also goes counter to our corpus measurements. We then argued that the two forces combined, with iconicity as a primary and frequency as a secondary factor, might explain a bit better the variety in the distribution of suppletion that we found in our sample.

We are fully aware that our investigation of the impact of frequency on suppletion leaves much to be desired. We have been able to consider only corpora for a small number of languages, and the ones that we have had access to are not uniform with respect to size or type of discourse, and other factors that probably are of relevance. Furthermore, we have been somewhat opportunistic in regard to our counting procedure, which ideally should be more in line with the most recent studies of lexical processing. Finally, the push-and-pull of the presence of a (partial) paradigm in a language should be added as an independent factor, rather than just as the lower limit of suppletion.

Nonetheless, we contend that even if all the above were to be catered for, frequency is unlikely to provide a comprehensive account of the distribution of suppletion in person paradigms. Nor is the interaction of iconicity and frequency likely to suffice. The structure of person paradigms is the result of a host of interacting factors and diachronic pressures, both language internal and external, which we do not yet fully understand. We trust, however that our investigation of suppletion has shed some light on the role of two of these.

Abbreviations

1 first person, 2 second person, 3 third person, abs absolutive, acc accusative, dat dative, du dual, erg ergative, ex exclusive, f feminine, in inclusive, m masculine, n neuter, nom nominative, pauc paucal, pl plural, sg singular

References

Bakker, Dik, André Müller, Viveka Velupillai, Søren Wichmann, Cecil Brown, Pamela Brown, Dmitry Egorov, Robert Mailhammer, Anthony Grant & Eric W. Holman. 2009. Adding typology to lexicostatistics: a combined approach to language classification. *Linguistic Typology* 13(1). 167–179.

Berkenfield, Catie. 2001. The role of frequency in the realization of English *that*. In Joan Bybee & Paul Hopper (eds.), 281–308.

Bhat, D. N. Shankara. 1967. Lexical suppletion in baby talk. *Anthropological Linguistics* 9(5). 33–36.

Bobaljik, Jonathan David. 2012. *Universals in Comparative Morphology*. Cambridge, MA: MIT Press.

Brown, Cecil H., Eric W. Holman, Søren Wichmann & Viveka Velupillai. (2008). Automated classification of the world's languages: A description of the method and prelimary results. *STUF* 61. 285–308.

Bybee, Joan. 1985. *Morphology. A study of the relation between meaning and form*. Amsterdam: John Benjamins.

Bybee, Joan. 2005. Restrictions on phonemes in affixes: A crosslinguistic test of a popular hypothesis. *Linguistic Typology* 9(2). 165–222.

Bybee, Joan. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Bybee, Joan & Joanne Scheibman. 1999. The effect of usage on degrees of constituency: the reduction of *don't* in English. *Linguistics* 37. 575–596.

Bybee, Joan & Paul Hopper (eds.). 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.

Cinque, Guglielmo. 1999. *Adverbs and Functional heads. A Crosslinguistic Perspective*. New York: Oxford University Press.

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell.

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

Corbett, Greville. 2005. Suppletion in personal pronouns: theory vs. practice and the place of reproducibility in typology. Linguistic Typology 9(1). 1–23.

Corbett, Greville. 2007. Canonical typology, suppletion, and possible words. Language 83(1). 8–42.

Corbett, Greville, Andrew Hippisley, Dunstan Brown & Paul Marriott. 2001. Frequency, regularity and the paradigm: A perspective from Russian on a complex relation. In Joan Bybee and Paul Hopper (eds.), 201–226.

Croft, William. 2003. *Typology and Universals*. Cambridge: Cambridge University Press.

Daniel, Michael. 2011. Plurality in Independent Personal Pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 35. (http://wals.info/chapter/35; accessed on 2013-02-24).

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

Dressler, Wolfgang. 1990. Sketching submorphemes within natural morphology. In Julian Mendez Dosuna & Carmen Pensado (eds.), 33–41.

Dressler, Wolfgang & Lavina Merlini Barbaresi. 1994. *Morphopragmatics. Diminutives and Intensifiers in Italian, German, and Other Languages*. Berlin & New York: De Gruyter Mouton.

Dryer, Matthew S. 1989. Large linguistic areas and language sampling. *Studies in Language* 13. 257–292.

Dryer, Matthew S. & Martin Haspelmath (eds.). 2011. *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. (Available online at http://wals.info)

Fertig, David. 1998. Suppletion, natural morphology, and diagrammaticity. *Linguistics* 36(6). 1065–1091.

Fidelholtz, James L. 1975. Word frequency and vowel reduction in English. *CLS* 11. 200–214.

Gordon, Raymond G., Jr. (ed.). 2005. *Ethnologue*. 15th Edition. SIL International. (www.ethnologue.com)

Haiman, John. 1985a. *Natural Syntax*. Cambridge: Cambridge University Press.

Haiman, John. 1985b. *Iconicity in Syntax*. Amsterdam: John Benjamins.

Hampe, Beate & Christian Lehmann. 2013. Partial coreference. This volume.

Harnisch, Rüdiger. 1990. Morphologische Irregularität – Gebrauchshäufigkeit – psychische Nähe. Ein Zusammenhang im empirischen Befund und in seiner theoretischen Tragweite. In Julian Mendez Dosuna & Carmen Pensado (eds.), 53–64.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42(1). 1–46.

Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19(1): 1–33.

Hollmann, Willem & Anna Siewierska. (2007). A construction grammar account of possessive constructions in Lancashire dialect: some advantages and challenges. *English Language and Linguistics* 11(2). 407–424.

Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In W. Christie (ed.), *Current progress in historical linguistics*, 96–105. Amsterdam: North Holland.

Juge, Matthew L. 1999. On the rise of suppletion in verbal paradigms. *Berkeley Linguistic Society* 25. 183–194.

Levenshtein, Vladimir I. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Cybernetics and Control Theory* 10. 707–710.

Lloyd, Richard. 1989. Bound and Minor Words in Baruya. *Data papers in Papua New Guinea Languages Volume* 35. Summer Institute of Linguistics.

Lyons, John. 1968. *Introduction to Theoretical Linguistics*. London: Cambridge University Press.

Magnuson, James S., James S. Dixon, Michael K. Tanenhaus & Richard N. Aslin. 2007. The Dynamics of Lexical Competition During Spoken Word Recognition. *Cognitive Science* 34. 1–24.

Maiden, Martin. 1992. Irregularity as a determinant of morphological change. *Journal of Linguistics* 28. 285–312.

Maiden, Martin. 2004. When lexemes become allomorphs– On the genesis of suppletion. *Folia Linguistica* 38(3–4). 227–256.

Markey, Tom. 1985. On suppletion. *Diachronica* 2. 51–66.

Mayerthaler, Willi. 1981. *Morphologische Natürlichkeit*. (Linguistische Forschungen 28.) Wiesbaden: Athenaion.

McQueen, James M., Dennis Norris & Ann Cutler. (1994). Competition in Spoken Word Recognition: Spotting Words in Other Words. *Journal of Experimental psychology: Learning, Memory and Cognition* 20(3). 621–638.

Mel'čuk, Igor. 1994. Suppletion: toward a logical analysis of the concept. *Studies In Language* 18(2). 339–410.

Mendez Dosuna, Julian & Carmen Pensado (eds.). *Naturalists at Krems* (Papers from the workshop on Natural Phonology and Natural Morphology, Krems, 1–7 July 1988). Salamanca: University of Salamanca.

Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.

Moravcsik, Edith. 1994. Group plural, associative plural or cohort plural. Email document LINGUIST List vol 5-681. 11 June 1994 ISSN 1068-4875.

Moravcsik, Edith. 2003. A semantic analysis of associative plurals. *Studies in Language* 27(3). 469–504.

Murthy, B. Lalitha & K. V. Subbarao. 2000. Lexical anaphors and pronouns in Mizo. In Barbara C. Lust, Kashi Wali, James W. Gair & K. V. Subbarao (eds), *Lexical Anaphors and Pronouns in Selected South Asian Languages*, 776–835. Berlin & New York: De Gruyter Mouton.

Nichols, Johanna. (2013). The origin and evolution of case-suppletive pronouns: Eurasian evidence. This volume.

Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago and London: The University of Chicago Press.

Nida, Eugene A. 1963. The identification of morphemes. In Martin Joos (ed.), *Readings in Linguistics*, 3rd ed., 255–271. New York: American Council of Learned Societies. (Reprinted from Language 24 [1948], 414–431).

Nübling, Damaris. 2000. *Prinzipen der Irregularisierung*. Tübingen: Niemeyer.
Owens, Jonathan. 1985. *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg: Buske.
Peirce, Charles. 1932. *Philosophical Writings, Vol. 2*. Cambridge: Harvard University Press.
Pike, Kenneth. 1965. Non-linear order and anti-redundancy in German morphological matrices. *Zeitschrift für Mundartforschung* 31. 193–221.
Riese, Timothy. 2001. *Vogul*. Munich: Lincom Europa.
Rijkhoff, Jan & Dik Bakker. 1998. Language sampling. *Linguistic Typology* 2(3). 263–314.
Ronneberger-Sibold, Elke. 1980. *Sprachverwendung– Sprachsystem: Ökonomie und Wandel* (Linguistische Arbeiten 87). Tübingen: Niemeyer.
Rudes, Blair. 1980. On the Nature of Verbal Suppletion. *Linguistics* 18. 655–676.
Schuchardt, Hugo. 1885. *Über die Lautgesetze: gegen die Junggrammatiker*. Berlin: R. Oppenheim.
Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.
Siewierska, Anna & Maria Papastathi. (2011). Third person plurals in the languages of
Europe: typological and methodological issues. *Linguistics* 49(3). 575–610.
Slater, Keith W. 2003. Mangghuer. In Juha Janhunen (ed.), *The Mongolic Languages,* 307–324. London & New York: Routledge.
Smith-Stark, T. Cedric. 1974. The plurality split. In Michael W. La Galy, Robert A. Fox and Ashton Bruck (eds.), *Papers from the Tenth Regional Meeting, Chicago Linguistic Society,* 657–671. Chicago: Chicago Linguistic Society.
Tomasello, Michael. 2003. *Constructing Language. A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
Tryon, Darrell T. 1970. *An Introduction to Maranungku (Northern Australia).* Canberra: Linguistic Circle of Canberra.
Urban, Greg. 1985. Ergativity and Accusativity in Shokleng (Gê). *International Journal of American Linguistics* 51(2). 164–187.
Van Valin, Robert D. & Randy J. LaPolla. 1997. *Syntax. Structure, meaning and function*. Cambridge: Cambridge University Press.
Veselinova, Ljuba. 2006. *Suppletion in verb paradigms: Bits and pieces of a puzzle*. Amsterdam: John Benjamins.
Voegelin, Charles F. & Florence M. Voegelin. 1977. *Classification and index of the world‘s languages* (Foundations of Linguistics Series). New York: Elsevier.
Werner, Otmar. 1987. Natürlichkeit und Nutzen morphologischer Irregularität. In Norbert Boretzky, Werner Enninger & Thomas Stolz (eds.), *Bochum – Essener Beiträge zur Sprachwandelforschung* IV, 289–316. Bochum: Brockmeyer.
Williamson, Kay. 1965. *A Grammar of the Kolokuma Dialect of Ijo*. Cambridge: Cambridge University Press.
Zipf, George K. 1965 [1935]. *The Psycho-Biology of Language: An Introduction to Dynamic Philology.* Cambridge, MA: The MIT Press.

Corpora

Language	Source	Web Location
English	BNC	http://www.natcorp.ox.ac.uk/
Estonian	MDC	http://www.cl.ut.ee/korpused/morfliides/

French	Beech	http://www.llas.ac.uk/resources/mb/80
German	DGD	http://dsav-oeff.ids-mannheim.de/
Hebrew	HSC	http://www.tau.ac.il/humanities/semitic/cosih.html
Italian	BADIP	http://badip.uni-graz.at/
Polish	Pelcra	http://nkjp.uni.lodz.pl/index.jsp
Portuguese	Davies	http://www.corpusdoportugues.org/
Russian	RNC	http://www.ruscorpora.ru/en/
Scots	SCOTS	http://www.scottishcorpus.ac.uk/
Spanish	Davies	http://www.corpusdelespanol.org/
Swedish	GSCL	http://www.ling.gu.se/projekt/tal/

Languages

Language Sample S488 (n=488)
NA = distinction in marking of Nom/Abs and Acc/Erg (n=178)
NN = no number marking (n=12)

Languages in italics are not part of the genealogically controlled subsample S350

Afro-Asiatic: *A.1*: Hausa(NA); *A.2*: Lele; *A.8:* Gude; *Arabic*: Arabic (Egyptian), Maltese; *Atlas:* Shilha,Tamazight; *Bole:* Bole; *Coptic:* Coptic; *Cushitic:* Angass; *Dizoid:* Dizi(NA); *Gimira: Gimira(NA); North:* Beja(NA), *Bilin(NA),* Geez; *Oromo:* Oromo (Harar) (NA); *Rendille-Boni: Boni; South: Ari(NA), Burunge,* Galila(NA), *Hamer(NA),* Iraqw, *Maale(NA); West: Basketo(NA), Hozo(NA)*

Alacalufan: Kawesqar(NN)

Algic: *Ojibwa*, Passamaquoddy-Maliseet, Wiyot, Yurok(NA)

Altaic: Daur(NA), *Even(NA)*, *Evenki(NA)*, *Khalka Mongolian(KA)*, *Mangghuer(NA)*, Turkish(NA), *Tuvin*, Udihe(NA), *Uyghur*

Andamanese: Onge(NA)

Araucanian: Mapuche

Arawakan: Apurinã, Arawak, Resígaro

Australian: Arabana(NA), Garawa(NA), Gooniyandi(NA), Gunya(NA), Kalkatungu(NA), Malakmalak, Maranungku, Martuthunira(NA), Maung, Nunggubuyu, Nyulnyul(NA), Pitjantjatjara(NA), Ungarinjin, Uradhi(NA), Wambaya, Wardaman(NA), Warlpiri(NA), Warrgamay

Austro-Asiatic: *Mon-Khmer*: Bugan, Cambodian, Car(NA), *Gorum(NA)*, Khasi, Khmu, Mon, Palyu, Ruc, Sedang, Semelai(NA), *Taoih*, *Temiar*; *Munda*: *Kera*, Mundari

Austronesian: *Adzera*: Adzera; *Aneityum*: *Anejom; Are*: *Gapapaiwa; Bali-Vitu*: *Bali-Vitu; Bariai*: *Kabana, Kove, Lusi; Central*: *Cemuhi, Kokota; Chamorro*: Chamorro; *Central Malayo Polynesian*: *Alune,* Arguni*, Buru,* Dawera-Daweloor, *Kisar, Leti,* Manggarai, Selaru(NA), *Sikka,* Tetun, Tugun, West Damar; *East*: Kele, Larike; *East Fijian*: *Fijian (Boumaa)(NA); East Makian-Gane*: Taba; *East Uvean-Niuafo'ou*: *Niuafoou; East Vanuatu*: *Ambrym (Southeast), Mwotlap, Raga; Erromanga*: Sye; *Futunic*: *Ifira-Mele; Gela*: *Gela; Ikiribati*: *Kiribatese; Jayapura Bay*: *Tobati; Kaili:* Uma; *Kairiru*: *Kairiru; Kilivila*: Kilivila; *Korap*: *Arop-Lokep; Labu*: *Labu; Lamenu-Lewo*: *Lamenu; Local Malay*: Indonesian;

Longgu: *Longgu; Loyalty Islands*: *Iai, Nengone; Malagasy*: Malagasy(NA); *Malekula Central*: *Vinmavis, Port Sandwich; Marquesic*: *Marquesan; Nimoa-Sudest*: *Sudest; North*: *Jabem, Sakao; Paiwanic*: Paiwan; *Palauan*: Palauan; *Pasismanua*: *Kaulong; Patpatar-Tolai*: *Siar; Piva-Banoni*: *Banoni; Rotuman*: *Rotuman; San Cristobal*: Arosi; *Saposa-Tinputz*: *Taiof; Sarmi*: *Sobei; St. Matthias*: Mussau; *Tagalog*: Tagalog(NA); *Tahitic*: Maori; *Trukic*: *Puluwat, Ulithian, Woleian; Unclassified*: Rejang(NA); *Utupua*: *Tanimbili; Vanikoro*: *Buma; West*: Roviana(NA); *West Fijian*: *Nadrog; West Santo*: *Tamabo; Western Malayo-Polynesian*: Cebuano, Ida'an(NA), Ilokano, Kapampangan, Ma'anyan, Manobo Cotabato, Muna, Nias(NA), Sama (Sinama), Sasak, Tboli; *Xaracuu-Xaragure*: *Xaracuu; Yapese*: Yapese; *Zire-Tiri*: *Tinrin*

Aymaran: Jaqaru

Barbacoan-Paezan: Awa Pit(NA), Tsafiki

Basque: Basque(NA)

Caddoan: Wichita

Cahuapanan: Chayahuita

Cariban: Carib(NN), Hixkaryana, Macushi

Chapacura-Wanham: Wari'

Chibchan: Bribri(NA), Guaymi(NA), Rama(NA)

Choco: Epena Pedee(NA)

Chon: Selknam, Tehuelche

Chukotko-Kamchatkan: Chukchi (Telqep)(NA)

Chumash: Chumash

Coahuiltecan: Tonkawa(NA)

Creole: Berbice Dutch, Kituba(NA), Mauritian Creole(NA), Ndyuka(NA), Palenquero, Tok Pisin

Dravidian: Brahui(NA), Kannada(NA), *Kodava(NA),* Malayalam(NA), Tamil(NA)

East Bird's Head: Sougb

East Papuan: Kuot, Nasioi(NA), Santa Cruz

Eskimo-Aleut: Yupik(NA)

Geelvink Bay: *Barapasi*, Saweru, Tarungare

Gulf: *Atakapa, Chitimacha*, Tunica

Hmong-Mien: Hmong Njua, Iu Mien

Hokan: Karok, Pomo (Southeastern)(NA), Washo

Huavean: Huave(NA)

Indo-European: Albanian(NA), Armenian (Eastern)(NA), Bengali(NA), *Bulgarian(NA)*, *Breton*, *Catalan(NA), Chali(NA), Croatian(NA)*, *Czech(NA), Danish(NA), French(NA), Gaelic,* Greek(NA), Gujarati(NA), Icelandic(NA), *Irish, Italian(NA)*, Kashmiri(NA), Latvian(NA), Lithuanian(NA), *Pashto(NA),* Polish(NA), *Portuguese(NA), Rumanian(NA),* Spanish(NA), Swedish(NA), Talysh (Northern)(NA), Welsh(NA)

Iroquoian: Cherokee(NN), Mohawk

Isolate: Ainu (Classical), Burmeso, Burushaski(NA), Candoshi(NA), Cayuvava(NA), Itonama, Jicaque(NA), Korean, Kutenai, Kwaza(NA), Mosetén, Movima, Nahali(NA), Nivkh (Gilyak), Porome, Puinave, Trumai(NA), Waorani, Warao(NA), Yaghan(NA), Yuchi, Yurakare, Zuni(NA)

Japanese: Japanese(NN)

Kartvelian: Georgian(NA)

Katukian: Kanamari

Keres: Acoma(NN)

Khoisan: *Ani*, Nama(NA)

Kiowa Tanoan: Kiowa(NN)

Kwomtari-Baibari: Momu

Macro-Ge: Bororo, Chiquitano, Kaingang, Karaja, Xokleng

Maku: Hupde(NA)

Mascoian: Lengua Mascoy

Mataco-Guaicuru: Toba

Mayan: Huastec, Jacaltec, Kekchi, Tzutujil

Misumalpan: Miskito

Mixe-Zoque: Oluta Popoluca, Zoque

Mura: Pirahã(NN)

Muskogean: Koasati

Na-Dene: Carrier, Haida(NA), Tlingit

Nambiquaran: Nambiquara

Niger-Congo: *Adamawa-Ubangi:* Bai; *Atlantic:* Izon, Wolof; *Bantoid: Befang, Lamnso,* Limbum(NA), Vute; *Busa:* Bokobaru; *Cangin:* Noon; *Central Niger Congo:* Konni(NA); *Dagaari:* Dagaare; *Dogon:* Dogon; *Dowayo:* Doyayo; *East:* Ibibio(NA); *Edekiri:* Yoruba(NA); *Edoid:* Edo(NA); *Idoid:* Eloyi(NA); *Igboid:* Igbo(NA); *Kainji:* Clela(NA); *Kissi:* Kisi(NA); *Koh:* Koh Lakka; *Kordofanian:* Katla, Krongo(NN), Tagoi; *Kwa:* Akan, Chumbarung, Ewe; *Kweni-Yaoure: Yaoure(NA); Liberian:* Grebo; *Mande:* Jalonke, Mandinka, *Sisiqa; Moba: Bimoba; Momo: Mundani(NA); Nupoid:* Gbari; *Plateau 1:* Doka; *Southeast:* Dagbani; *Swahili: Swahili; Ukaan:* Ukaan; *Unclassified:* Fali; *Western:* Godie; *Zande-Nzakara:* Zande

Nilo-Saharan: *Bagirmi:* Bagirmi; *Bari: Kuku; Berta:* Berta(NA); *Fur:* Fur(NA), *Kanuri: Dongolese Nubian(NA)*, Kanuri; *Komuz:* Kwama(NA); *Kunama:* Kunama; *Lango-Acholi:* Lango; *Lendu:* Ngiti; *Maba:* Mesalit(NA); *Murle:* Murle(NA); *Nandi: Nandi(NA); Ngangea-So:* So; *Songhai*: Songhai*; Teso: Teso; Turkana: Turkana(NA); Unclassified: Shabo(NA); Western:* Nuer

North Caucasian: Abkhaz, Chechen(NA), *Ingush(NA), Lak(NA),* Lezgian(NA)

Oto-Manguean: Copala Trique, Otomi, Popoloc Metzontla, Zapotec

Panoan: Marubo(NA), Matses(NA), Shipibo-Konibo(NA)

Peba-Yaguan: Yagua, *Yava*

Penutian: Nez Perce(NA), Siuslaw(NA), Takelma, Tsimshian, *Wintun(NA)*

Pidgin: Chinook Jargon

Quechuan: *Quechua Ayacucho(NA),* Quechua Huanuco(NA)

Salishan: Coeur d'Alene, Halkomelem, Lummi

Salivan: Piaroa

Sepik-Ramu: *Awtuw,* Chambri, Gapun(NA), Ngala, Rao(NA), Yessan Mayo(NA), *Yimas*

Sino-Tibetan: Chinese: Cantonese; Tibeto-Burman: Burmese(NA), Byangsi(NA), Chamling, Chepang(NA), Dulong, *Jinuo,* Kayah Li Eastern, Lepcha(NA), Lipo, *Lisu*, Lushai(NA), *Manchad(NA),* Qiang Southern, Tinan(NA), *Tinani(NA); Unclassified: Bawm, Nisu*, Pumi Northern

Siouan: Catawba, Lakhota, *Tutelo(NN)*

Skou: *I'saka,* Skou, *Vanimo*

Subitaba-Tlapanec: Tlapanec

Tacanan: *Araona(NA),* Cavineña(NA)

Tai-Kadai: Dong, Gelao, Thai

Tarascan: Tarascan

Torricelli: Au, Bukiyip, Olo, Walman
Totonacan: Totonac Misantla

Trans New Guinea: *Aldelbert Range*: *Mauwake(NA); Angan Proper*: *Kapau(NA); Baraic*: *Barai; Binanderean Proper*: *Suena; Brahman*: Tauya(NA); *Dumut*: Wambon; *Eastern:* Una(NA); *Eleman*: Kaki Ae(NA); Gum: Amele; Inland Gulf: Minanibai; *Kalam-Kobon*: Kobon(NA); *Kamano-Yagaria*: Hua(NA); *Koiaric*: *Koiali Mountain(NA); Kowan*: *Waskia(NA); Madang-Albert*: Kimaghama(NA); *Main Section*: Amanab(NA), Baruya, Binandere(NA); *Marind Proper*: *Marind;* Morwap: Elseng(NN); *Nimboran*: Nimboran(NN); *Numugenan*: *Usan; Oksapmin*: Oksapmin; *Senagi*: Kamberataro(NA), *Menggwa Dla(NN); South Bird's Head: Adang(NA); Teberan-Pawaian*: Folopa(NA); *Trans-Fly:* Kiwai Southern; *Turama-Kikorian:* Rumu

Tucanoan: Barasano(NA), Cubeo(NA), Retuarã

Tupian: Kanoe, Karo, Munduruku, Urubu-Kaapor(NA)

Unclassified: Birale, Yaruro(NA)

Uralic: Finnish(NA), Hungarian(NA), *Kamas(NA),* Nenets(NA), *Ostyak(NA), Udmurt(NA), Voghul Northern(NA)*

Uru-Chipaya: Chipaya, *Uru*

Uto-Aztecan: Comanche(NA), Kawaiisu(NA), Pipil, *Yaqui(NA)*

Wakashan: Nootka

West Papuan: Hatam, Maybrat, West Makian

Witotoan: Bora(NA), *Witoto*

Yanomam: Sanuma

Yeniseian: Ket

Yukaghir: Yukaghir(NA)

Yuki-Wappo: Wappo(NA)

Zaparoan: Iquito