J. Linguistics **47** (2011), 1–29. © Cambridge University Press 2010
doi:10.1017/S0022226710000022 First published online 6 April 2010

Defectiveness and homophony avoidance[1]

MATTHEW BAERMAN

Surrey Morphology Group, University of Surrey

(Received 7 October 2009; revised 22 February 2010)

The idea that certain morphological and phonological irregularities are due to speakers' desire to avoid homophony is widely invoked, yet has also come under strong criticism as an explanation which is neither necessary nor sufficient. In most cases there is no way to resolve the question, since the assumption that something is being avoided is itself a theoretical construct. In this article I attempt to address this last difficulty by looking at gaps in inflectional paradigms – where it is clear that SOMETHING is being avoided – that plausibly correlate with potential homophony. These fall into two types: (i) LEXICAL, where portions of the paradigms of two lexeme would be homophonous, and (ii) PARADIGMATIC (i.e. syncretism), where forms within the paradigm of a single lexeme would be homophonous. Case studies of Tuvaluan, Russian, Mazatec, Tamashek and Icelandic confirm the effects of homophony avoidance as a genuine, if non-deterministic, principle.

1. Homophony avoidance

The idea that certain morphological and phonological irregularities can be traced to a desire to avoid homophony has a long history, stretching (at least) from historical grammars of the nineteenth century (e.g. Diez 1838: 210)[2] through to the identity avoidance constraints of Optimality Theory (Crosswhite 1999, Rebrus & Törkenczy 2005). But homophony avoidance as an explanatory principle has also has some spirited detractors, most famously Lass (1980).

As an illustration of the principle, as well as the difficulties inherent in it, consider the paradigms in Table 1, from Teiwa, a Papuan language of Alor

[1] The research and writing of this paper were funded by the Arts and Humanities Research Council (UK) under grant number AH/D001579/1, and the European Research Council under grant number ERC-2008-AdG-230268 MORPHOLOGY. Their support is gratefully acknowledged. An earlier version of this paper was presented at the Décembrettes 6 conference in Bordeaux, 6 December 2008. Thanks to the audience there for their questions and comments. I would also like to thank Kristín Bjarnadóttir, Dunstan Brown, Patricia Cabredo-Hofherr, Carol Capen, Greville Corbett, Gunnar Ólafur Hansson, Jóhannes Gísli Jónsson and Maarten Kossmann, as well as two anonymous *JL* referees, for helpful discussion of various points.

[2] His example is the unusual development of the first and second person plural perfect in Romanian, which ought to have been homophonous with the present. No doubt earlier examples could be found.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	C-initial		V-initial			
	'hand'		'ear'		'heart'	
1SG	na	-tan	n	-uar wa'	n	-u'uk
1PL	ni	-tan	**ni**	-uar wa'	n	-u'uk
1PL INCL	pi	-tan	p	-uar wa'	p	-u'uk
2SG	ha	-tan	h	-uar wa'	h	-u'uk
2PL	yi	-tan	y	-uar wa'	y	-u'uk
3SG	ga	-tan	g	-uar wa'	g	-u'uk
3PL	gi	-tan	**gi**	-uar wa'	g	-u'uk
DISTRIBUTIVE	ta	-tan	t	-uar wa'	t	-u'uk

Table 1
Teiwa inalienably possessed nouns (Klamer 2010). Exceptional CV- prefixes with V-initial stems shown in boldface.

island, Indonesia (see Klamer 2010). Nouns may be prefixed for the person and number of the possessor. As a general rule, the shape of the prefix is CV- with C-initial noun stems, while with V-initial noun stems such as *-uar wa'* 'ear' the vowel of the prefix is deleted. But this vowel-deletion rule is not exceptionless: in the 1PL and 3PL, the vowel is retained. Now it happens that these are also the values for which deletion of the vowel would neutralize the distinction between singular and plural; otherwise, there is nothing phonological, morphological or semantic that sets apart these cells as a class. Therefore, we can say that there is a rule of vowel deletion which is blocked should it lead to homophony of singular and plural. But this irregularity only applies to some nouns; with others, such as *-u'uk* 'heart', vowel deletion applies without exception, leading to homophony.

Thus, while homophony avoidance may provide an explanation, it introduces an additional explanandum: when does it apply, and when not? Further, there is another morphological irregularity without any recoverable explanation that still needs to be accommodated: the sporadically attested alternative form *a-* of the third person prefix, e.g. *ga-tewar* or *a-tewar* 'his walking' (Klamer 2010).[3] In the light of such considerations, homophony avoidance seems neither sufficient nor necessary as an explanation of irregularities. And yet, slippery though the notion might be, there is nothing implausible in the idea that the maintenance of contrast plays SOME role in the organization of paradigms, and its appeal to linguists is evident by its longevity and frequent invocation.[4] Any resolution of the status of

[3] This matches the freestanding possessive pronoun.

[4] For a crude measure of its popularity, note that a search on Google Books (February 2009) just for the collocations "avoid homophony/homonymy", "éviter l'homophonie/homonymie" and "Vermeidung der Homophonie/Homonymie" yielded 804 hits.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

homophony avoidance is hindered by the fact that the thing which is being avoided is itself a theoretical construct. Thus, in Table 1, the expectation that there are homophonous forms is created by the assumption of a regular rule of vowel deletion. But of course, it could be that there is no such rule, and hence nothing to be avoided: the paradigms are simply irregular, and the grammar need tell us only how, not why.

There may, however, be a way around this difficulty. While language-specific diachronic and synchronic rules may be open to debate, we can safely assume that, barring semantic incompatibility, inflectional rules should produce a complete paradigm for every inflectable lexeme. Where we find instances of defectiveness, that is, gaps in inflectional paradigms, we can be fairly sure that SOMETHING is being avoided.[5] Of course, defectiveness as such may be attributable to a number of different causes, such as phonological or morphological constraints, alongside outright stipulation (Hetzron 1975, Fanselow & Féry 2002, Baerman & Corbett 2010). But, as I argue below, there are also examples of gaps that can convincingly be correlated with some expected homophony, and that this is evidence that homophony avoidance can play a decisive role in determining the shape of the paradigm. The paper is structured as follows. Section 2 deals with homophony that obtains between (forms of) different lexemes. Section 3 deals with paradigmatic homophony (i.e. syncretism), that obtains between different forms of the same lexeme. In Section 4 I conclude that there is evidence that homophony avoidance may play a role both in diachronic change and in synchronic rule interaction. I should note here that this research was originally undertaken to demonstrate the opposite point.

2. LEXICAL HOMOPHONY

I take it as relatively uncontroversial that we may find instances where an entire lexeme *Y* is avoided outright due to homophony with some other lexeme *X*, as occurs with taboo avoidance. But since we are looking for paradigm gaps, we need to find instances where avoidance affects only part of the paradigm. Probably the best-known examples come from French, and for that reason we ought to consider them straight off – if only to lay them aside again, as they are not entirely convincing. In the normative standard, the verb *clore* 'close' is missing considerable chunks of its paradigm, in particular the first and second person plural of the present, and all of the imperfect. According to some sources (e.g. Dauzat 1944: 116), this was

[5] The question of homophony avoidance as an explanation for gaps in DERIVATION has certainly been raised before (see Plank 1981: 166–170), without, as far as I can tell, a clear consensus having been reached. The problem is that while we normally expect inflectional operations to be applicable to any relevant lexeme (barring gross semantic incompatibility), we do not normally expect derivational operations to be quite so productive; gaps are part of the game. So the absence of a given derivational form is itself a dubious explanandum.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Middle French				Modern French	
	(a) clo(u)er 'nail'		(b) clore 'close'		(c) clore 'close'	
	SINGULAR	PLURAL	SINGULAR	PLURAL	SINGULAR	PLURAL
1	clo(u)e	clo(u)ons	clos	clouons	clos	–
2	clo(u)es	clo(u)ez	clos	clouez	clos	–
3	clo(u)e	clo(u)ent	clôt	closent	clôt	closent

Table 2
French (homophonous forms shaded).

a result of homophony clash with the corresponding forms of *clo*(*u*)*er* 'nail', as illustrated in Table 2).

But there is an alternative explanation. Gilliéron (1919) treats this as a case of conventional lexeme loss with subsequent repair. That is, the partial homophony of *clo*(*u*)*er* and *clore*, as well as a certain degree of semantic overlap ('nail (shut)' ≅ 'close'), led to the extinction of *clore*, which went out of use by the sixteenth century. However, the past participle of *clore*, namely *clos*, remained in adjectival use. The lexeme was then artificially resuscitated by the French Academy, which nevertheless limited itself to recommending those forms which could be deduced from the past participle, according to the usual distribution of stems in the paradigm (see e.g. Maiden 2005). Thus, according to Gilliéron, the pattern of homophony in middle French and the defectiveness of Modern French are unrelated products of the organization of stem alternations characteristic of French verbs; that is, the pattern in Table 2(c) is not the RESULT of the patterns in Table 2(a, b). Gilliéron also provides similar analyses for other putative examples of homophony-avoidance-induced defectiveness, such as *frire* 'fry' and *choir* 'fall'. In any event, the reconstruction is largely speculation, so the value of the French examples is uncertain.[6] But in the following sections I offer what seem to be clearer instances of homophony avoidance.

[6] Other examples that might be familiar to the reader suffer because the correlation is imprecise. In Russian, there are a number of verbs defective for their first person singular, which is an environment in which a regular stem alternation occurs. For a few verbs this alternation would make them homophonous with another lexeme, and some sources ascribe the defectiveness to homophony avoidance just for these verbs (Rozental′, Dzandzakova & Kabanova 2007: 427f.). But there are even more defective verbs where this potential homophony does not play a role, so some other explanation for the gaps must be found in any case. Albright (2003) cites the example of Spanish *abolir* 'abolish', which is missing its first singular indicative, which some speakers attribute to potential homophony with *abuelo* 'grandfather'. But the verb is missing other parts of its paradigm as well, where homophony avoidance is not an issue.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	SINGULAR	PLURAL, LOCATIVE
NEAR SPEAKER	tee-nei	ko-nei
NEAR ADDRESSEE	tee-naa	ko-naa
NEUTRAL	tee-laa	ko-laa

Table 3
Tuvaluan demonstrative/relative pronoun/adjective.

2.1 *Tuvaluan*

A fairly straightforward example of defectiveness due to lexical homophony avoidance comes from Tuvaluan (Austronesian, Oceanic). There is a demonstrative-cum-relative pronoun/adjective which is suppletive between the singular (*tee-*) and the plural/locative (*ko-*). It is inflected for a three-way deictic distinction as in Table 3.

However, as Besnier (2000: 419) observes

> A few older speakers use two archaic forms *kinei* and *kinaa* instead of the plural demonstrative adjectives and locative pronouns *konei* and *konaa*. There does not seem to be an archaic form for the third person, corresponding to *kolaa*. The third-person form would be homophonous with the word *kilaa* 'hairless', which has improprietous undertones, and it is very likely that the incomplete paradigm is due to word tabooing. (Attempts to elicit the missing form were invariably met with embarrassment or guffaws.)

This is a fairly classic taboo-motivated type of homophony avoidance, in this case limited to one form out of a larger paradigm rather than the entire lexeme. Presumably this occurs only because the root itself is obsolescent, and because a ready alternative exists. (Admittedly this does not answer the question of what prior generations did.)

2.2 *Russian*

The Russian indefinite pronouns *nékto* 'somebody, a certain' and *néčto* 'something' each have only one form: *nékto* has only a nominative, as in (1), while *néčto* has only a nominative/accusative, as in (2). (Most examples are taken from the Russian National Corpus (RNC); exact source details for examples (1)–(11) are provided in Appendix 1.)

(1) **Nékto** prixódit v kondíterskuju i trébuet sebé tórt …
somebody comes into café and asks.for self cake
'Somebody walks into a café and asks for a piece of cake …'

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Indefinite pronouns		Interrogative pronouns	
	'somebody, a certain'	'something'	'who'	'what'
NOM	někto	něčto	kto	čto
ACC	–	něčto	kogó	čto
GEN	–	–	kogó	čegó
LOC	–	–	kom	čëm
DAT	–	–	komú	čemú
INS	–	–	kem	čem

Table 4
Russian indefinite versus interrogative pronouns.

(2) (a) Verojátno, **něčto** podóbnoe proisxodílo …
probably something[NOM] similar happened
'Probably something similar happened …'
(b) A sočetánie daët **něčto** nóvoe.
and combination gives something[ACC] new
'And the combination yields something new.'

Compared to other semantically overlapping words, *někto* and *něčto* are infrequent and largely confined to a literary register; the RNC (queried in February 2009) records 1,902 instances of *někto* versus 25,165 of *któ-to* 'someone (specific)' and 6,903 of *kto-nibúd'* 'someone (non-specific)'. But their relative infrequency is no explanation for the lack of oblique case forms, particularly as the other semantically overlapping words have complete case paradigms. From a morphological point of view this restriction is also striking. The words are transparently composed of a prefix *ne-* plus the interrogative pronouns *kto* 'who' and *čto* 'what',[7] and these interrogative pronouns by themselves do decline, as shown in Table 4.

An interesting morphological correlation does emerge when we look at the paradigms of another pair of lexemes, the negative pronouns *nékogo* 'there is nobody' and *néčego* 'there is nothing'. These appear to be composed of exactly the same morphological elements: a prefix *ne-* plus the interrogative pronouns. But they are not used in the nominative (the genitive is conventionally used as the citation form), so their paradigms are at least roughly in complementary distribution with those of indefinite *někto*

[7] With the reservation that *čto* 'what' is pronounced [ʃ]*to*, while *něčto* 'something' is *ne*[tʃ]*to*. This distinction has a historical basis, the [tʃ] pronunciation being characteristic of Church Slavonic (see Section 2.2.2 below).

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Indefinite pronouns		Negative pronouns	
	animate: 'somebody, a certain'	inanimate: 'something'	animate: 'there's nobody'	inanimate: 'there's nothing'
NOM	někto	néčto	–	–
ACC	–	(P)néčto	ne(P)kogo	ne P čto
GEN	–	–	ne(P)kogo	ne(P)čego
LOC	–	–	ne P kom	ne P čem
DAT	–	–	ne(P)komu	ne(P)čemu
INS	–	–	ne(P)kem	ne(P)čem

Note: 'P' indicates the position of prepositions.

Table 5
Russian indefinite versus negative pronouns.

and *néčto* (Table 5), the only overlap being in the inanimate accusative. In fact, due to certain morphosyntactic details discussed below (section 2.2.1), in particular their contrasting behaviour with respect to prepositions, this overlap is only apparent, and the two sets of paradigms really are complementary.

As is shown below (Section 2.2.1), the fact that the negative pronouns lack a nominative has a clear syntactic motivation, so it is the defectiveness of the indefinite pronouns that needs to be accounted for. This is where homophony avoidance may come in as a possible explanation; thus, Zaliznjak (1973: 84, fn. 34) suggests that the oblique forms of indefinite *někto* and *néčto* are blocked because the corresponding forms have already been claimed by negative *nékogo* and *néčego*. Strictly speaking, such an analysis ought best to be regarded as poetic licence: the absence of oblique forms of *někto* and *néčto* is quite categorical, and there is none of the variation or uncertainty one would expect if the forms were being avoided. But there is evidence, admittedly indirect, that homophony avoidance in the past led to the state of affairs we have now. The story is a rather complicated one, requiring some background in the morphosyntactic and lexical peculiarities of the main players, as outlined below.

2.2.1 *Morphosyntax of* nékogo/néčego

The morphosyntax of the negative pronouns *nékogo* and *néčego* is such that they are excluded from the nominative entirely, and also from the inanimate accusative in most contexts. The lack of a nominative comes from the fact that they are used in subjectless sentences involving the copula

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

(which is null in the present tense) plus an infinitive.[8] Otherwise, case behaviour is normal, as illustrated by the various inflected forms of *nékogo* in examples (3)–(6): genitive/accusative in (3), locative in (4), dative in (5) and instrumental in (6).

(3) Króme sebjá obvinjátʹ býlo **nékogo**
except self.GEN blame.INF was.NEUT.SG nobody.ACC/GEN
'There was nobody to blame except himself.'

(4) **Ne** o **kom** bólʹše zabótitʹsja
no- about -body.LOC more worry.INF
'There's nobody to worry about any more.'

(5) Rasskazátʹ býlo **nékomu**
tell.INF was.NEUT.SG nobody.DAT
'There was nobody to tell it to.'

(6) a zamenítʹ ix búdet **nékem**
and replace.INF them.ACC will.be nobody.INS
'and there will be nobody to replace them with'

What would normally be a nominative subject in a finite clause may be expressed by the dative, as is typical with infinitives in Russian:

(7) Króme menjá tút dvérʹ otkrýtʹ **nékomu**
besides me here door[ACC] open.INF nobody.DAT
'There's nobody here to open the door besides me.'

The use of *nékogo* and *néčego* in the accusative is also highly restricted, but the reasons for this are more general in nature. Because the animate accusative is identical to the genitive in any case (see Table 4); this is a general property of most nominals in Russian), we need only consider *néčego* first. By a general rule, direct objects whose existence is negated outright – which certainly applies to these negative pronouns – take the genitive rather than accusative:

(8) (a) Ja vížu risk.
I see risk[ACC]
'I see a risk.'

(b) Ja ne vížu nikak-ogo risk-a.
I not see no-GEN risk-GEN
'I don't see any risk.'

This leaves only one context where the accusative might appear, namely when governed by a preposition that takes the accusative. This is then the one area where the paradigms of the indefinite pronouns *nékto*/*nečto* and the negative pronouns *nékogo*/*néčego* overlap. But significantly, because of their contrasting behaviour with prepositions, they do not coincide

[8] And as such they are not to be confused with the negative pronouns *niktó* 'nobody' and *ničegó* 'nothing', which are used in finite clauses.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

morphologically. The indefinite pronoun is treated as an indivisible whole, with the preposition preceding, as in (9), while with the negative pronoun, the preposition intervenes, as in (10).[9]

(9) Tý že sám, navérnoe, prošël **čerez néčto**
you EMPH self probably went through something[ACC]
podóbnoe …
similar
'You yourself probably went through something similar …'

(10) Nádo stroít′ nóvye, ináče èksportírovat′ néft′
necessary build.INF new.PL otherwise export.INF oil
búdet **né čerez čtó**
will no- through -thing[ACC]
'New ones need to be built, otherwise there will be nothing to export the oil through.'

As a consequence of these various factors, there is no overlap between the paradigms of the indefinite pronouns *nékto*/*néčto* and the negative pronouns *nékogo*/*néčego*.

2.2.2 *Diachrony*

Earlier texts show that the indefinite pronouns *nékto* and *néčto* used to have non-nominative forms, as in (11), where the verb *iskat′* 'look for' governs the genitive.

(11) uvíděl ón mnóžestvo žénščin, kotóryja **něčego**
saw he multitude women.GEN.PL who.NOM.PL something.GEN
iskáli s velíkim priležániem.
looked.for.PL with great.INS diligence.INS
'he saw a multitude of women who were looking for something with great diligence' (*Zadig* 109)

On this basis we might propose that these forms were lost over time; in particular, that they were driven out in order to avoid homophony with the negative pronouns *nékogo*/*néčego*. This is probably the wrong way of looking at it. The indefinite pronouns are in fact borrowed lexemes, and I will suggest

[9] Originally, this construction involved a negated copula used in conjunction with the interrogative pronoun (used as a free relative):

Older construction					*Modern Russian equivalent*			
ně	bě	emu	**kogo**	poslati	emu	bylo	**nekogo**	poslat′
not	was	him.DAT	who.ACC	send.INF	him.DAT	was	nobody.ACC	send.INF

'there was no one he could send' (Galician-Volynian Chronicle, 13th c.)

The reattachment of the negator *ne-* to the pronoun was probably a result of the later reduction of the present tense copula to null. The contemporary treatment of prepositions with *nékogo*/*néčego* is a reflection of this former independence of *ne*.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

that the role homophony avoidance played was as a constraint on borrowing. To make this clear, some background on the history of Russian is in order.

The confrontation of the indefinite and negative pronouns is ultimately an aspect of the confrontation between the two constituent strains of literary Russian, Church Slavonic versus the vernacular. The history of the literary language starts around the tenth century with the adoption of a literary idiom, Church Slavonic, itself originally based on a variety of Balkan Slavonic (presumably that spoken around Thessalonica). Although the differentiation of the various Slavonic languages from each other was at this point still in its early stages, Church Slavonic and vernacular Russian will nonetheless have differed in a number of important points. The subsequent history of the Russian literary language involves a continual interplay between the two, with influences going in both directions. The literary language as finally codified in the 18th century is something of a hybrid, vernacular Russian with countless Church Slavonic elements incorporated into it.

In the case at hand, the indefinite pronouns *někto* and *néčto* were originally Church Slavonic, while the negative pronouns *nékogo* and *néčego* were vernacular Russian. The indefinite pronouns form part of a small class of *wh*-words with the Church Slavonic prefix *ne-* 'some…' (homophonous with the ordinary sentence negator *ne* 'not'),[10] e.g. *kotóryj* 'who/which (relative pronoun)' ∼ *někotoryj* 'some, certain'; *kogdá* 'when' ∼ *nékogda* 'sometime'. In Church Slavonic, *někto* and *néčto* were fully inflected. However, there is no evidence that their oblique forms were ever successfully incorporated into anything that might be described as literary Russian.

Unfortunately, for a number of reasons it is difficult to offer much in the way of concrete evidence that this was really the case. First, the fluidity of the relationship between Church Slavonic and the vernacular means that one can never be too certain about identifying a given text as representing the donor or recipient idiom, especially in earlier periods. In the case of *néčto*, the evidence provided by Unbegaun (1935: 386) and Cocron (1962: 173) suggests that it entered the vernacular-based written norm in the 17th century (they do not discuss the possibility of oblique forms). The 18th-century texts available online in the Russian National Corpus and the *Russkaja virtual'naja biblioteka* (Russian Virtual Library) yield a grand total of one example

[10] Ultimately, the two are related, with the indefinite derived from the negative by lengthening of the vowel (Trubačev 1997: 92f.). The Church Slavonic reflex of the originally lengthened vowel was represented by a distinct grapheme *ě*, thus *ně* 'some' versus *ne* 'not'. In Russian the two have merged, with a purely graphic distinction maintained until the spelling reforms of the early 20th century eliminated *ě*. Note that prior to the standardization of spelling in the 18th century the distinction was inconsistent (e.g. see fn. 9 with *ně* for expected *ne*). Note also that Church Slavonic lacked the negative pronouns *nékogo* and *néčego* (having retained the older construction illustrated in fn. 9) so the homophonic clash under discussion would not have arisen.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	FEMININE	MASCULINE	NEUTER	PLURAL
NOM	nék-aja	nék-ij	nék-oe	nék-ie
ACC	nék-uju	NOM/GEN	nék-oe	NOM/GEN
GEN	néko-ej	néko-ego	néko-ego	néki-ix
LOC	néko-ej	néko-em	néko-em	néki-ix
DAT	néko-ej	néko-emu	néko-emu	néki-im
INS	néko-ej(u)	néki-im	néki-im	néki-imi

Table 6
Archaic paradigm of Russian *nékij* 'certain', with oblique stem forms highlighted (where NOM/GEN variation is indicated, the form is determined by animacy; animates take the GEN form).

of an oblique form. Examples such as (11) above are vanishingly rare. All this is congruent with the notion that the emergent modern literary standard simply by-passed the oblique forms.

In the case of *nékto*, the *Dictionary of 18th-century Russian* reports that it had only a nominative form, which would seem to leave us nothing to work with. However, there is another lexeme, the adjective *nékij* '(a) certain', some of whose forms correspond to the expected oblique forms of *nékto*.[11] It turns out we can fruitfully use *nékij* as a morphological surrogate for investigating the history of *nékto*. Before doing this, some explanation of its morphological peculiarities is necessary. *Nékij* is an adjective, and like all adjectives has a large paradigm, with distinct gender and number forms. What interests us is the behaviour of the oblique forms. As adopted from Church Slavonic, it had what is for Russian a quite unusual paradigm, shown in Table 6. While a normal Russian adjective has a single invariant

[11] Functionally, the two overlap when the referent is an animate singular, but neither subsumes the other. In the contemporary language their functions are distributed as follows:

		nékto	*nékij*
independent pronoun	'somebody'	√	
attributive w/ name	'a certain X'	√	√
attributive w/ common noun	'a certain person'	√	

Thus, when modifying a name, *nékto* and *nékij* are equivalent, but otherwise they diverge; in particular, *nékto* may function as an independent pronoun, while *nékij* is always an attributive modifier. But in Church Slavonic their functions were not clearly segregated. Given the formal overlap described in the main text below, examples of forms such as *nékogo*, *nékom*, *nékomu* in older texts are ambiguous: should they be construed as forms of *nékto* or of *nékij*? Presumably the authors of the *Dictionary of 18th-century Russian* have interpreted them all as forms of *nékij*, and so concluded that *nékto* fully lacked oblique forms in the period. This of course is perfectly possible, but the data are undeniably ambiguous.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

		Archaic pattern	Innovative pattern	Compare with *nékogo* 'nobody'
FEM SG GEN/LOC/DAT/INS	A	néko-ej(u)	nék-oj(u)	
MASC/NEUT SG GEN	B	néko-ego	nék-ogo	ne(P)kogo
MASC/NEUT SG LOC		néko-em	nék-om	ne P kom
MASC/NEUT SG DAT		néko-emu	nék-omu	ne(P)komu
MASC/NEUT SG INS	C	néki-im	nék-im	ne(P)kem
PL GEN/LOC		néki-ix	nék-ix	
PL DAT		néki-im	nék-im	
PL INS		néki-imi	nék-imi	

Note: 'P' indicates the position of prepositions.

Table 7
Variation in the oblique forms of Russian *nékij* 'certain'; 'A', 'B' and 'C' forms highlighted (see text and Figure 1; note that the alternation between *e* and *o* in the inflectional ending is morphophonologically regular).

stem, *nékij* had three: *nék-* for the nominative/accusative, and *néko-* and *néki-* for the oblique cases.

Unsurprisingly, the tendency over time has been to adapt the paradigm of *nékij* to that of a normal adjective, eliminating the alternation in favour of a single invariant stem, based on the nominative/accusative. One consequence of levelling out the stem alternation is that some of these forms (the dative, genitive and locative of the masculine/neuter singular) should end up identical to the corresponding but missing forms of *nékto*, as shown in Table 7, where the paradigm is divided into three chunks. 'C' corresponds simply to *néki-*, while 'A' and 'B' correspond to *néko-*: 'A' where levelling would not bring about homophony with *nékogo*, and 'B' where it would. Now if it is true that the oblique forms of *nékto* were rejected due to homophony with *nékogo*, we should equally expect stem levelling of *nékij* to have been rejected where homophony *nékogo* would have resulted. And this appears to be not only what happened, but what continues to happen, judging by the data from the RNC, which contains texts from 1700 on (see also Graudina, Ickovič & Katlinskaja 1976: 250).

Figure 1 shows the developments over the last three centuries. Levelling of *néki-* ('C') was already well under way in the 18th century, and by the late 20th century it had disappeared completely. The behaviour of the stem *néko-* differs sharply according to its potential for homophony with *nékogo*. Where there is no danger of homophony ('A'), levelling continues over time, till by the beginning of the 21st century the innovative non-alternating stem accounts for more than 60% of the tokens. Where homophony is involved

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

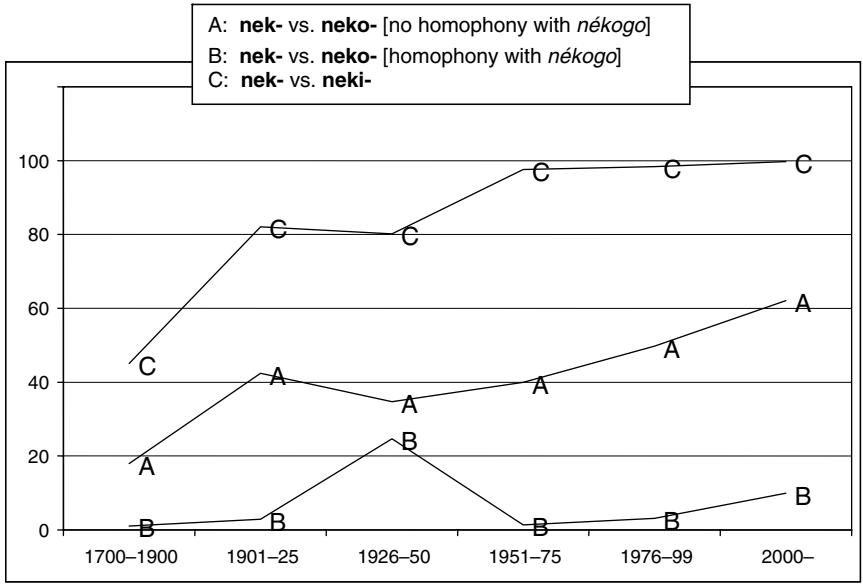


Figure 1
Percentage of innovative (non-alternating) oblique stem forms, per Russian National Corpus (queried in October 2008); raw data are given in Appendix 2.

('B'), levelling has been resisted, so that in the beginning of the 21st century levelled forms still account for only around 10% of the tokens.

Thus, although the historical record does not tell us exactly what happened to the oblique forms of *nékto*, the behaviour of *nékij* suggests that there is something about the forms themselves that speakers reject. And while there is no direct evidence that homophony avoidance is responsible, nothing else plausibly explains the distinction between 'A' and 'B' forms seen in Figure 1. Admittedly, the correlation is not perfect: there is no potential homophony in the locative, since the locative is only ever used with a preposition, resulting in a distinct construction (thus *o néko*(*e*)*m*… 'about a certain …' versus *ne o kom* '(there's) nobody about whom'), and yet levelling of locative *nekoem* to *nekom* appears nonetheless to be resisted to the same extent as with the other 'B' forms (see Appendix 2 for precise figures). I leave it to the reader to decide how serious a problem this is for the proposal I have offered. Overall, it does appear to me that the behaviour of *nékij* suggests that the negative pronoun *nékogo* is able to exert a repellent force on potentially homophonous forms of other lexemes, and that this provides indirect evidence that homophony avoidance determined which forms of the indefinite pronouns *nékto* and *néčto* were ultimately borrowed into Russian from Church Slavonic.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Neutral aspect		Incompletive aspect	
	AFFIRMATIVE	NEGATIVE	AFFIRMATIVE	NEGATIVE
1SG	ba^{3}nẽh^{31}	ba^{2}nẽh^{21}	kua^{3}nẽh^{31}	kua^{2}nẽh^{21}
2SG	ča^{3}nĩh^{31}	ča^{2}nĩh^{21}	ča^{4}nĩh^{41}	–
3	ba^{3}nĩh^{31}	ba^{2}nĩh^{21}	kua^{4}nĩh^{41}	–
1INCL	ča^{3}nẽh^{31}	ča^{2}nẽh^{21}	ča^{4}nẽh^{41}	–
1PL	ča^{3}nĩh^{314}	ča^{2}nĩh^{214}	ča^{4}nĩh^{414}	–
2PL	ča^{3}nũh^{31}	ča^{2}nũh^{21}	ča^{4}nũh^{41}	–

Table 8
Chiquihuitlan Mazatec 'carry' (Jamieson 1982: 166).

3. Paradigmatic homophony (syncretism)

Just as lexical homophony is common across languages, homophony between different inflectional forms of the same lexeme (or 'syncretism') is nothing exceptional (see Baerman, Brown & Corbett 2005). And still, some defectiveness appears to be motivated by avoidance of this homophony.

3.1 *Chiquihuitlán Mazatec*

In Chiquihuitlán Mazatec, verbs inflect for subject person-number, aspect (neutral versus incompletive) and polarity (affirmative/negative). Jamieson (1982) reports that the transitive verb *ba^{3}nĩh^{31}* 'carry' is defective for all the forms of the incompletive negative except for the 1SG (see Table 8; in the orthography employed here, *Vh* indicates laryngealized vowel, tone is indicated by superscript numerals, ranging from '1' (high) to '4' (low), and nasalization is indicated by a tilde). There is no apparent semantic reason for this gap. The speaker consulted reported that forms of the (non-defective) intransitive verb *ya?a* 'carry' are used instead (Carol Capen, p.c.).

This is quite a striking gap: semantically, it is incoherent, nor does it correspond to any obvious morphological parameter, such as a stem alternation. I suggest it can be interpreted as an instance of homophony avoidance, specifically between the affirmative and negative: the verb *ba^{3}nĩh^{31}* 'carry' has phonological properties that result in the neutralization of polarity distinctions in the incompletive for all values but 1SG.

Chiquihuitlán Mazatec inflection involves a complex interaction of morphological processes, so this is not immediately apparent from glancing at the paradigm. Negation is marked partly by (i) a distinct ending, which Jamieson represents as an underlying *-ĩ*, but which in most instances merges with the subject marker, and partly by (ii) tonal alternations, which varies

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Neutral aspect		Incompletive aspect	
	AFFIRMATIVE	NEGATIVE	AFFIRMATIVE	NEGATIVE
1SG	ka^{3}ntæ^{1}	ka^{2}ntẽ^{21}	ska^{3}ntæ^{1}	ska^{2}ntẽ^{21}
2SG	ča^{3}nti^{1}	ča^{2}ntĩ^{21}	ča^{4}nti^{1}	ča^{4}ntĩ^{41}
3	ka^{3}nti^{1}	ka^{2}ntĩ^{21}	ska^{4}nti^{1}	ska^{4}ntĩ^{41}
1INCL	ča^{3}ntẽ^{31}	ča^{2}ntẽ^{21}	ča^{4}ntẽ^{41}	ča^{4}ntẽ^{41}
1PL	ča^{3}ntĩ^{14}	ča^{2}ntĩ^{214}	ča^{4}ntĩ^{14}	ča^{4}ntĩ^{414}
2PL	ča^{3}ntũ^{1}	ča^{2}ntũ^{21}	ča^{4}ntũ^{1}	ča^{4}ntũ^{41}

Table 9
Chiquihuitlan Mazatec 'throw out' (Jamieson 1982: 166–167).

according to the accentual class of the verb.[12] These two systems are illustrated in Table 9 by the verb *ka^{3}nti^{1}* 'throw away', which belongs to the same conjugation class[13] and tonal class (class A) as *ba^{3}nĩh^{31}*.

The negative ending here is realized solely as nasalization of the person-number ending (note that by a regular phonological rule, *-æ* and *-e* merge under nasalization). Since nasalization is already a secondary feature of the plural endings, negation is contrastively marked by nasalization only in the singular. The tonal alternations are given in (12).[14] Note that the 1INCL incompletive affirmative in Table 9 remains unchanged, because it fails to meet the conditions for the tonal alternation: it has no tone 3 to raise, while the extension of the tone of the penultimate syllable onto the final is already a general characteristic of the 1INCL.

(12) Affirmative →negative tone alternations (in tone class A)
  (a) any tone 3 in the last two syllables is raised to 2
  (b) the tone of the penultimate syllable is extended onto the final (e.g. syll^{4} syll^{1}→syll^{4} syll^{41})

However, the verb *ba^{3}nĩh^{31}* neutralizes many of the affirmative ∼ negative contrasts on account of its phonological characteristics. First, the stem ends in a nasal consonant, and by a regular phonological rule, vowels immediately following a nasal consonant are nasalized. Since the negative endings are distinguished from the affirmative solely by nasalization, the endings themselves fail to mark negation. Second, the final syllable has a laryngealized

[12] There are four major tonal classes (A–D) which are in turn divided into further subtypes.

[13] The *i*-final verbs (the conjugation classes are based on the stem-final vowel, which in Jamieson's analysis merges with the (vocalic) subject marker that follows).

[14] Here and elsewhere Jamieson states the tone rules differently, but I believe the present formulation yields the same result, but more succinctly, at least for the purpose at hand.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Neutral aspect		Incompletive aspect	
	AFFIRMATIVE	NEGATIVE	AFFIRMATIVE	NEGATIVE
1SG	ci^{3}ka^{3}nẽh^{31}	ci^{3}ka^{2}nẽh^{21}	ci^{4}ka^{3}nẽh^{31}	ci^{4}ka^{2}nẽh^{21}
2SG	nĩ^{3}ka^{3}nĩh^{31}	nĩ^{3}ka^{2}nĩh^{21}	nĩ^{4}ka^{3}nĩh^{31}	nĩ^{4}ka^{2}nĩh^{21}
3	ci^{3}ka^{3}nĩh^{31}	ci^{3}ka^{2}nĩh^{21}	ci^{4}ka^{3}nĩh^{31}	ci^{4}ka^{2}nĩh^{21}
1INCL	nĩ^{3}ka^{3}nẽh^{31}	nĩ^{3}ka^{2}nẽh^{21}	nĩ^{4}ka^{3}nẽh^{31}	nĩ^{4}ka^{2}nẽh^{21}
1PL	nĩ^{3}ka^{3}nĩh^{314}	nĩ^{3}ka^{2}nĩh^{214}	nĩ^{4}ka^{3}nĩh^{314}	nĩ^{4}ka^{2}nĩh^{214}
2PL	nĩ^{3}ka^{3}nũh^{31}	nĩ^{3}ka^{2}nũh^{21}	nĩ^{4}ka^{3}nũh^{31}	nĩ^{4}ka^{3}nũh^{21}

Table 10
Chiquihuitlan Mazatec 'load' (Jamieson 1982: 166).

vowel, which in this accentual class attracts the tone of the preceding syllable (Jamieson 1982: 147), and so the tonal spread described in (12b) is neutralized. Both these neutralizations are clearly visible in the neutral aspect paradigm in Table 8. As a result, the only way a nasalized, laryngealized stem such as *ba^{3}nĩh^{31}* has to mark negation is the tonal alternation of 3 to 2 in the last two syllables (12a). This is indeed how negation is realized in the neutral aspect. But in the accentual class that *ba^{3}nĩh^{31}* belongs to, the incompletive affirmative is characterized by a lowering of an initial tone 3 to 4 in all the non-1SG forms, as seen in both Table 8 and 9. Therefore the conditions for 3→2 tone lowering are not met outside the 1SG, and there are no morphological means to distinguish the negative from the affirmative in the non-1SG incompletive forms. This, I would suggest, accounts for the speaker's reluctance, as the negative forms would be homophonous with the affirmative. As further evidence of this, note that there is a prefixed causative formed from the same root, *ci^{3}ka^{3}nĩh^{31}* 'load' (i.e. 'cause to carry'), illustrated in Table 10, which undergoes the same neutralizations associated with nasalized, laryngealized roots, and so can only mark negation by the tone raising rule in (12a). And although it belongs to the same accentual class as *ba^{3}nĩh^{31}*, it is a three syllable verb. That means that the lowering of the initial tone 3 to 4 in the non-1SG incompletive leaves the penultimate syllable unaffected, which is then free to undergo negative tone raising (and, by (12b), the final syllable follows suit). The negative forms are thus distinct from the affirmative, and the paradigm is not defective.

One question that this account does not resolve is the status of the 1INCL incompletive of other verbs of this tonal class; as noted above, and illustrated in Table 9, this form does not distinguish affirmative and negative. Jamieson (1982: 165–167) lists such forms in paradigms given in her article. If the explanation above is correct, why are these paradigms not defective too? I can only speculate; one possibility may be that it is a by-product of elicitation conditions, i.e. the infelicity of affirmative/negative homophony

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Normal affixal system		
	short imperfective (V-initial)	long imperfective (C-initial)	perfective (C-initial)
1SG	ikwɑl-æɣ	tikwɑl-æɣ	kæwɑl-æɣ
2SG	t-ikwɑl-æd	tikwɑl-æd	kæwɑl-æd
3SG.MASC	ikwɑl	i-tikwɑl	kæwɑl
3SG.FEM	t-ikwɑl	tikwɑl	kæwɑl
1PL	n-ikwɑl	nə-tikwɑl	??
2PL.MASC	t-ikwɑl-æm	tikwɑl-æm	kæwɑl-æm
2PL.FEM	t-ikwɑl-mæt	tikwɑl-mæt	kæwɑl-mæt
3PL.MASC	ikwɑl-æn	tikwɑl-æn	kæwɑl-æn
3PL.FEM	ikwɑl-ænt	tikwɑl-ænt	kæwɑl-ænt

Table 11
Tamashek adjectival verb ('be black'), showing the normal set of vowel- and consonant-initial prefixes, alongside the prefixless perfective paradigm (Heath 2005: 388, 431f., 437).

only struck the speaker when larger chunks of the paradigm were being offered.[15]

3.2 *Tamashek*

Finite verbs in Tamashek (a Berber language of the Tuareg group) are inflected for person, number and gender of the subject through a combination of prefixes and suffixes. There is a single set of suffixes, while the prefixal paradigm depends on the stem-initial segment (vowel-initial versus consonant-initial). The contrast is shown in the first two columns in Table 11: the vowel-initial set has a prefix *t-* (second person and the third singular feminine) which the consonant-initial set lacks, while the consonant-initial set has a prefix *i-* (third singular masculine) which the vowel-initial set lacks. The opposition of vowel- and consonant-initial forms is in turn tied to the system of stem alternations: the short imperfective, the long imperfective and the perfective.[16] As a general rule, the short imperfective

[15] Note also that there is another, less frequently used means of negating verbs, involving a separate negator *ʔa⁴kũĩ⁴¹* (the negative of a verb meaning 'to be') used before the affirmative verb form (Jamieson 1982: 165; Capen 1996: 35). Whether its use might be favoured under these conditions is also unclear; the one grammar of Chiquihuitlán Mazatec (Jamieson 1988) contains no mention of this negation strategy.

[16] These terms, as well as 'adjectival verb' which follows, are what Heath (2005) uses. The terms vary widely in the literature on Berber languages; of particular note is that the adjectival verbs are often called 'statives'.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

and the perfective are vowel-initial and the long imperfective is consonant-initial. There is one important exception though, namely the class of verbs that Heath (2005) terms 'adjectival verbs', such as *kæwal-* 'be black'. Their perfective stem is consonant-initial, but in place of the set of prefixes normally found with consonant-initial stems, they have no prefix at all (see the third column of Table 11). And what is significant for our purposes, the first person plural is defective.

Since the subject inflection of the perfective of adjectival verbs is characterized by complete omission of prefixes, by the logic of this system, the third person singular and the first person plural should then simply be the bare stem. But, as Heath notes, speakers reject the first person plural interpretation of bare stem:

> Instead, a circumlocution or a specialized construction was offered to express senses like 'we became black' A T-ka (Timbuktu area, Kal Ansar) informant offered kæwɑl-áte-næɣ, a difficult-to-segment morphological oddity that seems to involve an apparent **preposition-like-extension** -áte- that takes the 1Pl suffix -næɣ. There are other prepositions ending in -e- before pronominal suffix [sic], so one could segment kæwɑl-á́t-e-næɣ but the only -æt suffix that can occur in such a position is FeSg Participle suffix -æt, so the construction is obscure. Another T-ka speaker, the R (Rharous area) speaker, offered a **circumlocution** with Reslt[ative] -ámós- 'be, become' and a plural relative clause: n-ámós [i kæwɑ́l-nen] 'we have become black ones'. (Heath 2005: 437f.)

This does look like a fairly straightforward instance of homophony avoidance, assuming blind application of a prefix-deleting rule. But the unexplained alternative form in *-æ-te-næɣ* that Heath cites suggests that there may be other rules lurking beneath the surface. This impression is confirmed when one considers the broader comparative Berber perspective: the perfective paradigm of the adjectival verbs is the product of a series of small-scale analogical developments, and the role of homophony avoidance is best understood in this diachronic context (for an overview of the full range of variation in this class, see Kossmann 2009).

Following Prasse (1973: 10f.) I assume that in Proto-Berber the normal system of verbal affixes and the perfective paradigm of adjectival verbs were completely distinct from each other. This state of affairs is still found in Nafusi Berber (Beguinot 1942): the normal system corresponds closely to that found in Tamashek,[17] while the perfective paradigm of adjectival verbs marks gender and number only (see Table 12).

[17] Third singular masculine *y-* is realized as *i-* before consonant-initial stems; this is also true of all subsequent tables. The omission of *t-* before consonant-intial stems seen in Table 11 appears to be restricted to Tamashek and, to a limited extent, the Tuareg dialect of Hoggar (Prasse 1973: 9).

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Normal system	Perfective adjectival verbs
1SG	… -eɣ	
2SG	t- … -ed	… (M)
3SG.MASC	y- …	… -yet (F)
3SG.FEM	t- …	
1PL	n- …	
2PL.MASC	t- … -em	
2PL.FEM	t- … -met	… -et
3PL.MASC	… -en	
3PL.FEM	… -net	

Table 12
Nafusi (Beguinot 1942: 45, 66–67). Forms that conflate person distinctions indicated by shading.

	Normal system	Perfective adjectival verbs
1SG	… -eɣ	… **-eɣ**
2SG	t- … -eḍ	… **-eḍ**
3SG.MASC	y- …	…
3SG.FEM	t- …	… -et
1PL	n- …	
2PL.MASC	t- … -em	
2PL.FEM	t- … -met	… -it
3PL.MASC	… -en	
3PL.FEM	… -net	

Table 13
Kabyle (Naït-Zerrad 1994: 16, 23). Forms that conflate person distinctions indicated by shading. Affixes borrowed from the normal system indicated by boldface type.

In other Berber languages, the normal system of affixes has been extended to greater or lesser degree to the perfective adjectival verb paradigm. Kabyle (Naït-Zerrad 1994) illustrates the minimal adaptation: only overt suffixes have been borrowed, and only in the singular (see Table 13). As a consequence, the original gender-number suffixes are now restricted to 3SG.

In Zenaga (Table 14), the same strategy is optionally applied to the plural,[18] with the overt suffixes (second and third person) taken from the normal verbal system, leaving the original plural suffix restricted to the first person.

[18] Some verbs show a preference for one or the other suffixal series. In particular, verbs that normally have human subjects tend towards the innovative series.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Normal system	Perfective adjectival verbs	
		Older	∼ innovative
1SG	… -äg	… **-äg**	… **-äg**
2SG	t- … -äḏ	… **-äḏ**	… **-äḏ**
3SG.MASC	y- …	…	…
3SG.FEM	t- …	… -äḏ	… -äḏ
1PL	n- …		… -iḏ
2PL.MASC	t- … -äm		… **-äm**
2PL.FEM	t- … -ämñaḏ	… -iḏ	… **-əmñäḏ**
3PL.MASC	… -än		… **-än**
3PL.FEM	… -əñ(ñäḏ)		… **-əñ(ñäḏ)**

Table 14
Zenaga (Taine-Cheikh 2003: 664–665). Forms that conflate person distinctions indicated by shading. Affixes borrowed from the normal system indicated by boldface type.

In the various Tamashek dialects, the adaptation of the suffixal system has proceeded one step further, embracing the null suffixes as well. All share the adoption of zero suffixation of the 3SG FEM, but differ in the treatment of the 1PL.[19] Recall that Heath (2005) describes two variants, either the one with the mystery ending *-æte-næɣ* ('variant 1' in Table 15), or the form is avoided altogether ('variant 2' in Table 15). The dialect of Adrar has both zero suffixation and the prefix *n-* (final column in Table 15).

I suggest the following scenario to account for this variation, illustrated in (Table 16). The first three reconstructed stages correspond to Nafusi, Kabyle and Zenaga, described above, leading to the adoption of the overt suffixes from the normal verbal system into the paradigm of the perfective adjectival verbs. Stage IV represents a hypothetical stage in which zero suffixation was borrowed in the singular, leaving the 1PL as the sole relic of the older system. (There is no direct evidence for this stage, but a parallel singular ∼ plural split can be found in some varieties of Tamajak (another language of the Turareg group), in which the whole paradigm has been normalized except for the 3PL; Prasse, Alojaly & Mohamed 2003: 422.) With this as a point of departure, the attested non-defective varieties of Tamashek represent attempts to normalize the 1PL:

- Addition of the 1PL suffix *-næɣ*, yielding Heath's mystery ending *æt-e-næɣ*. This *-næɣ* is drawn from a set of personal suffixes used with prepositions,

[19] In Sudlow's (2001) grammar of the dialects of Tamashek spoken in Burkina Faso, the section devoted to this class of verbs states simply that they differ from normal verbs in that they do not take prefixes, but the only forms actually cited are third person (p. 184f.), so it is not entirely certain how the 1PL behaves here.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Heath (2005), variant 1	Heath (2005), variant 2	Adrar
1SG	… -æɣ	… -æɣ	… -æɣ
2SG	… -æd	… -æd	… -æd
3SG.MASC	… -Ø	… -Ø	… -Ø
3SG.FEM	… -Ø	… -Ø	… -Ø
1PL	… -æte-næɣ	–	n- … -Ø
2PL.MASC	… -æm	… -æm	… -æm
2PL.FEM	… -mæt	… -mæt	… -mæt
3PL.MASC	… -æn	… -æn	… -æn
3PL.FEM	… -næt	… -næt	… -næt

Table 15
Variant 1PL forms for perfective adjectival verbs in Tamashek (Heath 2005 and Prasse & ăgg-Ălbosṭan ăg-Sidiyăn 1985: 24).

which are distinct from those used with nouns. (The status of the *-e* is uncertain, but there is at least one preposition that adds an *-e* in its plural forms: *jer* 'between ~ *jere-næɣ* 'between us' (Heath 2005: 290), so there is perhaps an analogy here.)[20] The result is a form which is overtly 1PL, but still anomalous in the context of the verbal system.

- The logical next step in the accommodation of this class to the normal verbal pattern would be adoption of zero suffixation in the 1PL, which would yield 3SG/1PL homophony. In place of this, Adrar has gone one step further, disambiguating the homophony through the adoption of the 1PL prefix as well. (Note that 3SG masculine prefix *i-* is not borrowed, so this clearly does not represent a wholesale adoption of the normal system of prefixes and suffixes, but rather is limited to the problematic 1PL.)

In this scenario, the defective paradigm described by Heath represents the transitional stage between the system in which morphological traces of the old paradigm are still retained (in the form of the ending *-ætenæɣ*), and the system found in Adrar, where the accommodation to the normal system has moved beyond the suffixes and embraced the prefixes as well. Homophony avoidance plays a crucial role in this account, but it is not the sole motivation: the gap in the paradigm results not just from rejection of the 3SG/1PL homophony, but from rejections of all the available morphological options.[21]

[20] One potential difficulty with this interpretation is that, when suffixed to a preposition, 1PL *-næɣ* normally assigns stress to the immediately preceding vowel (Heath 2005: 274).

[21] As with all the other examples discussed in this paper, the claim that speakers have rejected a particular homophony does not constitute a claim that rejection is inevitable. Thus, Cortade (1969: 41) describes the system in Tamahak, another language of the Tuareg group, as having exactly the homophony which Heath's Tamashek speakers appear to have rejected.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Normal verbal system	Perfective adjectival verbs					
		reconstructed				attested Tamashek	
		stage I	stage II	stage III	stage IV	(Heath 2005)	Adrar
1SG	… -æɣ	… (MASC) … -yæt (FEM)	… **-æɣ**	… **-æɣ**	… **-æɣ**	… **-æɣ**	… **-æɣ**
2SG	t- … -æd		… **-æd**	… **-æd**	… **-æd**	… **-æd**	… **-æd**
3SG.MASC	y- …		…	…	…	…	…
3SG.FEM	…		… -yæt	… -yæt	… **-Ø**	… **-Ø**	… **-Ø**
1PL	n- …	… -æt		… -æt	… -æt	… -ætenæɣ	**n-**… **-Ø**
2PL.MASC	t- … -æm			… **-æm**	… **-æm**	… **-æm**	… **-æm**
2PL.FEM	t- … -mæt		… -æt	… **-mæt**	… **-mæt**	… **-mæt**	… **-mæt**
3PL.MASC	… -æn			… **-æn**	… **-æn**	… **-æn**	… **-æn**
3PL.FEM	… -næt			… **-næt**	… **-næt**	… **-næt**	… **-næt**

stage I corresponds to Nafusi (Table12)
stage II corresponds to Kabyle (Table 13)
stage III corresponds to Zenaga (Table 14)
stage IV is a hypothetical intermediary

Table 16

Proposed diachrony for perfective adjectival verbs in Tamashek (non-defective paradigms). Forms that conflate person distinctions indicated by shading. Affixes borrowed from the normal system indicated by boldface type.

	MASCULINE		FEMININE			NEUTER
	'time'	farmer'	'lie'	'hole'	'tongue'	'eye'
NOM SG	tím-i	bónd-i	lyg-i	hol-a	tung-a	aug-a
ACC SG	tím-a	bónd-a	lyg-i	hol-u	tung-u	aug-a
DAT SG	tím-a	bónd-a	lyg-i	hol-u	tung-u	aug-a
GEN SG	tím-a	bónd-a	lyg-i	hol-u	tung-u	aug-a
NOM PL	tím-ar	bænd-ur	lyg-ar	hol-ur	tung-ur	aug-u
ACC PL	tím-a	bænd-ur	lyg-ar	hol-ur	tung-ur	aug-u
DAT PL	tím-um	bænd-um	lyg-um	hol-um	tung-um	aug-um
GEN PL	tím-a	bænd-a	lyg-a	??	tung-na	aug-na

Table 17
Icelandic weak noun classes.

3.3 *Icelandic*

Icelandic has two genitive plural endings, *-a* and *-na*. For some nouns neither is felicitous, with homophony avoidance being a possible partial explanation (Jóhannes Gísli Jónsson, Gunnar Ólafur Hansson, p.c.). The ending *-a* is clearly the default, being found with all classes of strong nouns and most classes of weak nouns.[22] The ending *-na* is highly restricted: it is found with the weak neuters (a tiny class with about a dozen members; Thomson 1987: 234) and some weak feminines (Thomson 1987: 142 estimates that they number around 75). The various weak noun paradigm classes are shown in Table 17. It is among the weak feminines whose nominative singular ends in *-a* that difficulties arise.

Those such as *tunga*, which have the lexically specified GEN PL *-na*, are unproblematic. But there are others, such as *hola*, which for some speakers decidedly do not have the *-na* ending, but the default *-a* ending is equally bad. That homophony avoidance may be playing a role is suggested by the fact that the addition of the definite suffix to *hola* makes it considerably better (Gunnar Ólafur Hansson, p.c.), since in the definite paradigm the NOM SG/GEN PL homophony is disambiguated (see Table 18).

It should also be borne in mind that historically, homophony avoidance looks as if it might have played a role in the distribution of *-na*. Originally (that is, in Proto-Scandinavian), *-na* was the ending for all weak nouns. It has since yielded to *-a* precisely in that class where it is never homophonous

[22] From a comparative Germanic perspective, declension classes are divided into strong and weak. The weak classes were originally characterized by an extended stem in *-Vn* in the oblique cases of the singular, and in the plural.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

	Indefinite		Definite	
	SINGULAR	PLURAL	SINGULAR	PLURAL
NOM	hol-a	hol-ur	hol-an	hol-urnar
ACC	hol-u	hol-ur	hol-una	hol-urnar
DAT	hol-u	hol-um	hol-unni	hol-unum
GEN	hol-u	hol-a??	hol-unnar	hol-anna(?)

Table 18
Icelandic feminine noun 'hole'.

with the NOM SG, namely the masculines, and retained where it is always homophonous, namely the neuters. The feminines remain a mixed bag; the available evidence shows a great deal of variation among speakers, with few if any generalizations that would help determine which ending a noun ought to follow (Bjarnadóttir 2007). It is likely that this uncertainty plays a role in making the genitive plural form problematic.

I will not even try to speculate why this particular homophony is so obnoxious, given the rampant syncretism which otherwise characterizes Icelandic nominal paradigms. I limit myself to pointing out that this is the only class of nouns in which the nominative singular is syncretic with some other case in the plural (except for instances where the NOM SG is syncretic with the NOM PL, and this in turn is syncretic with some other plural case). To sum up, this is an intriguing and understudied phenomenon; aside from a brief mention in Jónsson (1927: 16), I am aware of no published work specifically addressing the defectiveness of words of this class. In particular, corpus work may well reveal something interesting, as the words so affected are sufficiently frequent that we may find observable effects. However, such work awaits the production of a disambiguated corpus.

4. Conclusion

The case studies presented in this paper demonstrate that homophony avoidance may play a role in defectiveness, and consequently that homophony avoidance is a principle that can play a role in grammar in general. Whether this principle can be incorporated into a formal theory of grammar is quite another question. There are no solid cross-linguistic generalizations about what constitutes intolerable homophony. Most linguists would probably accept the idea that taboo avoidance, as in Tuvaluan (Section 2.1), is a real phenomenon, but we can hardly characterize TABOO in any general way. And the neutralization of the contrast of affirmative and negative, as in Chiquihuitlán Mazatec (Section 3.1), certainly constitutes about as

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

pernicious an example of homophony as one could wish for,[23] but most putative examples involve neutralizations which are otherwise cross-linguistically common or even found elsewhere within the language itself (as in Teiwa). And yet there does appear to be an observable effect; even more, a remarkably persistent one, if the analysis of Russian offered above (Section 2.2) is correct.

In a recent study, Mondon (2009) proposes that the only useful constraint we can impose on homophony avoidance is on what it yields: it should be viewed as a condition affecting the choice among variants which are already available. This certainly seems plausible,[24] but what sort of status does that lend the principle? Another look at Tamashek (Section 3.2) may put it in the proper perspective. In the reconstruction offered above, defectiveness results from the rejection of both of two alternative morphological forms, where homophony avoidance only plays a role with one of them, while the other is objectionable for entirely parochial, paradigm-specific reasons (something on the order of the 'system congruity' of Natural Morphology; Wurzel 1984). This suggests it is a mistake to view homophony avoidance as universal functional principle (however seldom invoked). The term 'homophony' as such may lend itself to a universal characterization, just as terms such as 'prefixation' or 'suffixation' may, but whether or not it is avoided is no more generalizable across languages than whether, say, prefixed or suffixed forms are avoided.

[23] All the same, *could care less* in place of *couldn't care less* continues to spread in American English.

[24] This is in the spirit of Garret's (2008) proposal that paradigm levelling always involves a switch to a different, but already existing, inflectional class (one which lacks the offending alternation), rather than a direct manipulation of the paradigm.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

APPENDIX 1

Sources for the Russian examples in the text (all except for (8) and (11) were taken from the Russian National Corpus)

(1) M. V. Musijčuk. 2003. O sxodstve priemov ostroumija i mexanizmov postroenija paradoksal′nyx zadač. *Voprosy psixologii* 6.
(2a) Vasil′ Bykov. 1998. *Bednye ljudi.*
(2b) (no author given). 2004. Narodnyj kostjum: arxaika ili sovremennost′? *Narodnoe tvorčestvo* 1.
(3) Grigor Apojan. 2003. Nairi: triumf i drama. *Lebed′* (December 18).
(4) (no author given). 2003. Pomogite nam vstretit′sja. *Sel′skaja nov′* (October 7)
(5) Ljudmila Ulickaja. 2000. *Putešestvie v sed′muju storonu sveta.*
(6) (no author given). 2004. Žizn′, muzyka, studenty. *Ural′skij avtomobil′* (February 12)
(7) Andrej Gelasimov. 2001. *Foks Malder poxož na svin′ju.*
(8) [constructed example]
(9) Oleg Divov. 1998. *Molodye i sil′nye vyživut.*
(10) Julija Latynina. 2003. Mark Tven i ènergitičeskaja strategija. *Eženedel′nyj žurnal* (May 27).
(11) I. L. Goleniščevyj-Kutuzovyj. 1765. Translation of Voltaire's *Zadig*. Citation taken from the *Dictionary of 18th-century Russian* (entry under *néčto*).

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

APPENDIX 2

Stem variation in the oblique case forms of *nékij*: Raw numbers from the Russian National Corpus (queried in October 2008)

	Archaic/innovative	1700–1900	1901–1925	1926–1950	1951–1975	1976–1999	2000–
FEM SG	nékoej/nékoj	24/6	38/27	113/60	57/38	180/178	257/420
FEM SG INS	nékoeju/nékoju	8/1	0/1	0/0	0/0	0/0	0/0
MASC/NEUT GEN SG	nékoego/nékogo	118/0	73/3	121/36	92/2	335/11	592/64
MASC/NEUT LOC SG	nékoem/nékom	22/1	12/0	22/15	26/0	94/5	141/24
MASC/NEUT DAT SG	nékoemu/nékomu	39/1	14/0	22/3	26/0	66/0	112/5
MASC/NEUT INS SG DAT PL	nékiim(nékoim)/nékim	20(4)/14	5(1)/26	11/69	2/62	5/206	1(1)/410
GEN/LOC PL	nékiix(nékoix)/nékix	11(3)/7	0(1)/4	1/13	0/16	0/79	0/248
INS PL	nékiimi(nékoimi)/nékimi	1/1	0/2	0/7	0/3	0/19	0/78

Note: In earlier periods the feminine locative singular had the distinct ending *-oju*, which eventually yielded to the syncretic ending *-oj*.

REFERENCES

Albright, Adam. 2003. A quantitative study of Spanish paradigm gaps. *WCCFL* 22, 1–14.

Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2005. *The syntax–morphology interface: A study of syncretism* (Cambridge Studies in Linguistics 109). Cambridge: Cambridge University Press.

Baerman, Matthew & Greville G. Corbett. 2010. Defectiveness: Typology and diachrony. In Matthew Baerman, Greville G. Corbett & Dunstan Brown (eds.), *Defective paradigms: Missing forms and what they tell us* (Proceedings of the British Academy 163), 1–18. Oxford: The British Academy & Oxford University Press.

Beguinot, Francesco. 1942. *Il Berbero Nefûsi di Fassâṭo*. Rome: Instituto per l'oriente.

Besnier, Niko. 2000. *Tuvaluan: A Polynesian language of the Central Pacific*. London: Routledge.

Bjarnadóttir, Kristín. 2007. Enn um -na: um eignarfall fleirtölu af veikum kvenkynsnafnorðum í Beygingarlýsingu íslensks nútímamáls. Presented at 21st Rask Conference, University of Iceland, Reykjavík, January 21.

Capen, Carol Jamieson. 1996. *Diccionario mazateco de Chiquihuitlán, Oaxaca*. Tucson, AZ: Instituto Lingüístico de Verano.

Cocron, Friedrich. 1962. *La langue russe dans la seconde moitié du XVIIe siècle*. Paris: Institut d'études slaves.

Cortade, Jean-Marie. 1969. *Essai de grammaire touareg* (*dialecte de l'Ahaggar*). Algiers: Université d'Alger, Institut de Recherches Sahariennes.

Crosswhite, Katherine. 1999. Intra-paradigmatic homophony avoidance in two dialects of Slavic. In Matthew Gordon (ed.), *UCLA Working Papers in Linguistics* 1 (Papers in Phonology 2), 48–67. Los Angeles, CA: UCLA Department of Linguistics.

Dauzat, Albert. 1944. *La géographie linguistique*. Paris: Flammarion.

Diez, Friedrich. 1838. *Grammatik der romanischen Sprachen*, vol. 2. Bonn: Eduard Weber.

Fanselow, Gisbert & Caroline Féry. 2002. Ineffability in grammar. In Gisbert Fanselow & Caroline Féry (eds.), *Resolving conflicts in grammars: Optimality Theory in syntax, morphology, and phonology*, 265–307. Hamburg: Buske.

Garrett, Andrew. 2008. Paradigmatic uniformity and markedness. In Jeff Good (ed.), *Explaining linguistic universals: Historical convergence and universal grammar*, 125–143. Oxford: Oxford University Press.

Gilliéron, Jules. 1919. *La faillite de l'étymologie phonétique: Étude sur la défectivité des verbes*. Neuveville: Beerstecher.

Graudina, Ljudmila Karlovna, Viktor Aleksandrovič Ickovič & Lija Pavlovna Katlinskaja. 1976. *Grammatičeskaja pravil'nost' russskoj reči. Opyt častotno-stilističeskogo slovarja variantov*. Moscow: Nauka.

Heath, Jeffrey. 2005. *A grammar of Tamashek* (*Tuareg of Mali*). Berlin: Mouton de Gruyter.

Hetzron, Robert. 1975. Where the grammar fails. *Language* 51, 859–872.

Jamieson, Carol [Capen]. 1982. Conflated subsystems marking person and aspect in Chiquihuatlan Mazatec. *International Journal of American Linguistics* 48.2, 139–176.

Jamieson, Carol [Capen]. 1988. *Gramatica mazateca*. Mexico City: Instituto Lingüístico de Verano.

Jónsson, Snaebjörn. 1927. *A primer of modern Icelandic*. Oxford: Oxford University Press.

Klamer, Marian. 2010. *A grammar of Teiwa*. Berlin: De Gruyter Mouton.

Kossmann, Maarten. 2009. La flexion du prétérit d'état en berbère: questions de morphologie comparée. In Amina Mettouchi Chaker & Gérard Philipson (eds.), *Etudes de phonétique et linguistiques berbères. Hommage à Naïma Louali* (*1961–2005*), 147–167. Paris & Leuven: Peeters.

Lass, Roger. 1980. *On explaining language change*. Cambridge: Cambridge University Press.

Maiden, Martin. 2005. Morphological autonomy and diachrony. *Yearbook of Morphology* 2004, 137–175.

Mondon, Jean-François. 2009. *The nature of homophony and its effects on diachrony and synchrony*. Ph.D. dissertation, University of Pennsylvania.

Naït-Zerrad, Kamal. 1994. *Manuel de conjugaison kabyle*. Paris: L'Harmattan.

Plank, Frans. 1981. *Morphologische* (*Ir-*)*Regularitäten: Aspekte der Wortstrukturtheorie*. Tübingen: Narr.

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press

Prasse, Karl-G. 1973. *Manuel de grammaire touarègue*, vol. 6/7: *Verbe*. Copenhagen: Akademisk Forlag.
Prasse, Karl-G. & Ekhya ăgg-Ălbosṭan ăg-Sidiyăn. 1985. *Tableaux morphologiques: dialecte touareg de l'Adrar du Mali* (*berbère*). Copenhagen: Akademisk Forlag.
Prasse, Karl-G., Ghoubeïd Alojaly & Ghabdouane Mohamed. 2003. *Dictionnaire touareg–français* (*Niger*), vol. 2. Copenhagen: Museum Tusculanum Press.
Rebrus, Péter & Miklós Törkenczy. 2005. Uniformity and contrast in the Hungarian verbal paradigm. In Laura J. Downing, T. Alan Hall & Renate Raffelsiefen (eds.), *Paradigms in phonological theory*, 263–295. Oxford: Oxford University Press.
Rozental′, Ditmar E., Evgenija V. Dzandzakova & Natalija P. Kabanova. 2007. *Spravočnik po russkomu jazyku*, 5th edn. Moscow: Ajris-Press.
Russian National Corpus. http://www.ruscorpora.ru/ (accessed October 2008).
Russkaja virtual′naja biblioteka [Russian virtual library]. http://www.rvb.ru/ (accessed October 2008). [Includes a collection of major 18th c. texts.]
Slovar′ russkogo jazyka XVIII veka [Dictionary of 18th-century Russian], vol. 14. 2004. Saint Petersburg: Nauka.
Sudlow, David. 2001. *The Tamasheq of North-East Burkina Faso*. Cologne: Rüdiger Köppe.
Taine-Cheikh, Catherine. 2003. L'adjectif et la conjugaison suffixale en berbère. In Jérôme Lentin & Antoine Lonnet (eds.), *Mélanges David Cohen*, 661–674. Paris: Maisonneuve.
Thomson, Colin D. 1987. *Íslensk beygingafraeði*. Hamburg: Buske.
Trubačev, Oleg N. (ed.). 1997. *Ėtimologičeskij slovar′ slavjanskix jazykov*, vol. 24. Moscow: Nauka.
Unbegaun, Boris. 1935. *La langue russe au XVIe siècle* (*1500–1550*). Paris: Institut d'études slaves.
Wurzel, Wolfgang Ullrich. 1984. *Flexionsmorphologie und Natürlichkeit: ein Beitrag zur morphologischen Theoriebildung*. Berlin: Akademie.
Zaliznjak, Andrej A. 1973. O ponimanii termina 'padež' v lingvističeskix opisanijax. In Andrej A. Zaliznjak (ed.), *Problemy grammatičeskogo modelirovanija*, 53–87. Moscow: Nauka. [Also in Andrej A. Zaliznjak, *Russkoe imennoe slovoizmenenie*, 613–647. Moscow: Jazyki slavjanskoj kul′tury, 2002.]

Author's address: *Surrey Morphology Group, University of Surrey, Guildford, Surrey GU2 7XH, UK*
M.Baerman@surrey.ac.uk

https://doi.org/10.1017/S0022226710000022 Published online by Cambridge University Press