Contact-induced change in Dolgan : an investigation into the role of linguistic data for the reconstruction of a people's (pre)history

Stapert, E.L.

CHAPTER 4 LEXICON

4.1 INTRODUCTION

One linguistic domain in which contact influence can easily be recognised is the lexicon. Lexical items of foreign origin often stand out in their phonological or morphological structure and are therefore readily identifiable as copies from another language. However, this only applies to so-called full copies (see Section 3.3), where the complete unit of form and meaning are copied into the recipient language. Apart from these cases, there are many examples where either only the phonological form, or the meaning of a foreign lexical item is transferred to the recipient language. In the latter case this means that the phonological form of a word remains the same, while its semantic structure become more similar to the semantic structure of the source language. This phenomenon has been referred to by a variety of terms in the literature, including *Lehnbedeutung* 'semantic loan' (Blank 1997: 349, Betz 1949: 15), loanshift (Smith-Stark 1994: 17), lexical calquing (Ross 2007: 122)[1] or semantic borrowing (Geeraerts 2010: 29). Since it is impossible to identify this type of semantic transfer on the basis of the lexical form in the recipient language, the only way to detect changes of this kind is to make a fine-

[1] In earlier work Ross referred to this phenomenon as lexical metatypy (Ross 1996).

grained comparison of the semantic organisation of relevant semantic fields in the source language (SL) and the recipient language (RL).

In this chapter, lexical differences between Dolgan and Sakha are investigated, including both full copies and structural copies, with particular emphasis on instances of semantic change. The attested differences are analysed from a quantitative (4.4.1) as well as from a qualitative point of view (4.4.2) and in cases where a semantic difference between Dolgan and Sakha is attested it will be investigated whether its emergence could have been motivated by contact.

Since the lexicon is only one of the linguistic domains that need to be examined with respect to potential contact-induced change, the aim of this chapter is not to give an exhaustive inventory of differences across the entire lexicon. Rather, it focuses on a limited but comparable set of meanings in Dolgan and Sakha, allowing for an in-depth analysis of the quantitative as well as qualitative properties of the attested differences.

4.2 LEXICAL CHANGE AND LANGUAGE CONTACT

Most of the literature on contact-induced change in the lexicon focuses on the transfer of lexical copies from a source language into a recipient language. By default the tacit implication is that full copies are concerned, i.e. both phonological form as well as semantic structure are imported into the recipient language as a single unit. However, data from language contact studies show that this assumption is not always correct and it appears that the two aspects of De Saussure's linguistic sign can be, and often are, transferred separately. As mentioned above, the signified, corresponding to the semantics of a lexical item, may be copied independently from the signifier (or phonological form) into the recipient language, and can be projected onto existing native phonological forms, leading to a semantic change in this language. Although there is a comparatively small body of literature on this subdomain of lexical change and the sociolinguistic conditions underlying the phenomenon, it has been recognised and described by different scholars, and before presenting the data from Dolgan, a brief overview of relevant work in this domain is in order.

The categorisation of lexical change in most current theories is founded on work by Betz (1949) and Haugen (1950, 1953). Betz made a distinction between *Lehnwort* 'loanword' and *Lehnprägung* 'loan coinage' or 'loan meaning', whereby

Lehnwort refers to the transfer of both form and meaning (i.e. a full copy), and *Lehnprägung* to the transfer of meaning only (i.e. semantic copy). An example of the first category is the word *curry* in English. Here both the concept of 'spicy sauce' and the phonological form curry were adopted from Tamil *khari* 'sauce'. An example of the second category is the extension of the word *frio* 'cold temperature' in American Portuguese to include the additional meaning of 'cold infection' (Winford 2010: 172). This category also includes calques (or loan translations) of the type 'skyscraper', compounds of which the semantic structure has been copied into several languages, e.g. German *Wolkenkratzer*, Dutch *wolkenkrabber*, Russian *neboskreb*, French *gratte ciel*. This distinction was further refined by Haugen (1949: 289-290, 1950: 219-220), who subdivided Betz's category of *Lehnprägung* into loanshifts, including loan meanings (i.e. semantic copies) as illustrated above, and creations. A creation is defined as an innovative combination of native lexical items that is not directly copied from the source language, but is somehow inspired by the contact situation. An example is the Pima expression 'wrinkled buttocks' to refer to the concept of elephant that was foreign to the Pima speakers (Winford 2010: 173). In this case, the concept was introduced through contact, but the lexical form or semantic structure is not. Thus Haugen's loanword (i.e. full lexical copy) is defined as the copy of the unit of phonological form and semantic structure into a recipient language, whereas loan meanings (i.e. semantic copies) are "changes in the semantics of an RL word under influence from an SL word", (Winford 2010: 172). As mentioned above, later scholars have treated the latter phenomenon under the names of loanshift (Smith-Stark 1994: 17) or semantic borrowing (Geeraerts 2010: 29). The crucial difference between full lexical copies on the one hand and semantic copies on the other is that only for full lexical copies is there transfer of actual phonological form. For semantic copies (including purely semantic copies as well as calques) semantic structures are transferred from a source language into a recipient language, but these are mapped onto a native phonological form.

Most theoretical models of language contact differentiate between the copying of phonological form and the copying of grammatical or semantic structure (see Chapter 3). They make distinctions such as global vs. partial copying (Johanson 1992: 175), copying of matter vs. pattern (Matras and Sakel 2007: 829-830, 2008: 15), substance vs. schematic copies (Croft 2003a) or diffusion of form vs. diffusion of pattern (Aikhenvald 2003: 3). However, while some scholars would associate these different kinds of linguistic change with different socio-historical

and sociolinguistic conditions (e.g. Thomason and Kaufman 1988), different kinds of change within the lexical domain are rarely linked to different social conditions. One model that accounts for this distinction is the work by Ross in his discussion of metatypy, a term he uses to describe the restructuring of the recipient language on the model of a source language, whereby "the 'metatypised' (restructured) language maintains forms resembling those in its genetic relatives but the meanings of these forms have changed" (Ross 1996: 182). This restructuring may affect the grammar or morphology of the recipient language, but may equally well apply to the lexicon, in which case the phenomenon is referred to as lexical metatypy (Ross 1996: 191). To put it even stronger, Ross (2007: 124) seems to imply that morphosyntactic metatypy originates in lexical metatypy, which thus could be the source of more large-scaled restructuring of a language. He argues that semantic reorganisation develops out of different construals of reality that individuals are faced with while learning different languages, imposing an extra cognitive burden on the brain. Bilingual individuals "shed much of the burden by bringing the semantic organizations and reality construals of their two languages into line with each other" (Ross 1996: 204). They do that first by assimilating the ways of saying things in one language to the semantic organisation of their other language, which may then be extended to the restructuring of morphosyntactic categories.

Although Ross makes no explicit predictions as to whether this is the only setting in which lexical metatypy can occur, he describes it as a process that may occur in bilingual situations where a small emblematic language is modeled on a larger intergroup language (see Section 3.1.4.2). This social setting should be distinguished from that of the more frequent transfer of full lexical copies, which may enter a language in situations of casual contact and low degree of bilingualism among recipient language speakers. Thomason and Kaufman (1988: 77) predict that "with a minimum of cultural pressure we expect only lexical borrowing, and then only in non-basic vocabulary", which they define as culture-specific content words that are typically copied for cultural and/or functional reasons. They add that this type of transfer often happens in socially dominant languages that copy lexical items from numerically smaller groups. Winford (2010: 177) emphasises the role of 'need and prestige' for the occurrence of lexical copying, but does not mention any restrictions on the transfer of semantic structures.

4.3 METHODOLOGY AND CATEGORIES OF CLASSIFICATION

4.3.1 METHODOLOGY

As was described in detail in Section 1.2.2, the lexical comparison between Dolgan and Sakha was carried out on the basis of the set of 1,500 meanings from the Loanword Typology list. Although the purpose of the Loanword Typology project was narrower than the current aim to investigate lexical differences in general, the 1,500 item meaning list provides a useful starting point for comparing lexica in a structured way. This section provides the methodological background as well as the analytical framework that was used for the classification of lexical differences between Dolgan and Sakha.

While the registration of differences between Dolgan and Sakha is the first requirement for the establishment of lexical change, the identification alone does not provide any information on the direction of change. In order to arrive at a diachronic 'change' from the synchronic state implied by the word 'difference', it is necessary to identify a direction and an initiator of change. To achieve this, the attested differences between the Turkic languages Dolgan and Sakha were contrasted on the one hand with the linguistic encoding of these concepts in other Turkic languages and with Tungusic languages (in particular the contact language Evenki) on the other. This applies to differences in linguistic forms as well as to differences in their semantics. Through such a comparison with genealogically related languages on the one hand, and non-related but geographically neighbouring languages on the other it is possible to uncover typical semantic groupings within each language family, as well as deviations from these patterns, which in turn is crucial for the establishment of the direction of change (see also Section 1.2.3).

The general working hypotheses that I adhered to are the following:

1 a. If for a certain difference Sakha shows a feature which is typical for Turkic languages, but which is lacking in Dolgan, then it is more likely that the difference reflects a change in Dolgan.

b. If for a certain feature Dolgan conforms to a general Turkic pattern and Sakha does not, the innovation is more likely to have occurred in Sakha.

2 a. If for a certain difference Dolgan is similar to Evenki, and to Tungusic languages in general, contact with these languages could be an explanation for the change in Dolgan.

b. If the similarity exists just with Evenki and not with other Tungusic

languages, the possibility of contact-induced change in Evenki should be considered.

In practice this involved the following procedure. For each concept for which Dolgan differs from Sakha, a detailed comparison was made of the semantics of the lexical item in both languages based on data from fieldwork (from elicitation and narratives), complemented by extensive dictionaries for Dolgan (Stachowski 1993, 1998) Sakha (Sleptsov 1972, Pekarskij [1907 - 1930] 1958-1959, Voronkin 1995, 1999) and comparative dictionaries and grammars for Turkic (Sevortyan 1974, Erdal 1998, Tenishev 2001). Comparison with other Turkic languages then showed which language follows the general Turkic semantic pattern, and which language deviates from it. For example, a shared semantic grouping for Sakha and other Turkic languages versus a broader meaning in Dolgan could point to an extension of meaning in Dolgan. Conversely, an identical semantic pattern in Dolgan and other Turkic languages versus a different pattern in Sakha could reflect a change in Sakha. The hypothetical scenario where both Dolgan and Sakha differ from other Turkic languages in different ways was not attested and will therefore not be elaborated further. Comparison with Evenki (Boldyrev 1994, Myreeva 2004, Vasilievich 1968) and other Tungusic languages (Doerfer 2004, Benzing 1956), should then clarify whether contact with this language could explain the attested difference. For each concept it was investigated: a) whether the innovative pattern in Dolgan matches the pattern in Evenki; and b) whether Evenki displays a division of conceptual space found more generally within Tungusic, or whether it is an outlier within the family. The latter is important because if Evenki differs from other Tungusic languages, this could reflect a change in Evenki changed due to contact with Turkic languages. Taking all this information together: if Sakha follows the Turkic pattern and Evenki the Tungusic pattern, while the semantic organisation of Dolgan is different from Sakha in a way that matches Evenki, then there is reason to assume semantic transfer from Evenki to Dolgan. If, on the other hand, the different meaning in Dolgan does not match Evenki, the change in Dolgan could have developed as a result of language-internal motivations, such as language attrition. During this process, specific or infrequent words may fall into disuse, which may lead to the development of polysemy in the remaining lexical items.

4.3.2 CATEGORIES OF CLASSIFICATION

Any comparison can be conceived of as the description of one phenomenon with respect to another and involves a standard of comparison and a comparee. Without wanting to make a priori assumptions about the direction of change, for the purpose of terminology I have taken Sakha as the standard of comparison and have described the differences in Dolgan with respect to Sakha. The attested differences between Dolgan and Sakha were classified into six overarching categories that cover the various ways in which forms or meanings diverge in the two languages. These categories are represented in the first column of Table 4.1 and are further clarified below. For certain purposes, these six categories were further specified in the categories in the second column.

Table 4.1: Categories for classification of lexical differences between Dolgan and Sakha

Category	Subcategory
REPLACEMENT	- Russian copy
	- Evenki copy
	- Unknown
SEMANTIC CHANGE	- Broader
	- Narrower
	- Shift
CHANGE IN FORM	- Phonetic change
	- Morphological change
DESCRIPTIVE	- Descriptive phrase
	- Calque
ABSENCE	
?	

REPLACEMENT*:* for a concept, Dolgan and Sakha use different lexical forms and the form in Dolgan cannot be traced back to Sakha (or Turkic) origin. This category covers copies from Russian and Evenki, but also a few lexical items of unclear origin that could be copies from other languages or language-internal innovations in Dolgan. It is important to note that under the label 'Russian copy' two kinds of replacements are subsumed. On the one hand, this label covers cases where for a certain concept Sakha uses a Sakha word and Dolgan uses a word from Russian; on the other hand, it is applied to cases where the Russian copy in Dolgan is different

from the Russian copy that is used in Sakha. Russian copies shared by Dolgan and Sakha are not counted as differences between the two languages, and so are not considered here.

SEMANTIC CHANGE*:* this label is applied to cases where "[o]nly the meaning of a lexeme changes while other aspects of the item remain constant" (Wilkins 1996: 268). That is, a difference is classified as semantic change if for a concept Dolgan uses a lexical form that is also used in Sakha, but with a different meaning. The difference is classified as 'broader' when the Dolgan meaning covers a wider semantic space than it does in Sakha, as 'narrower' when it covers less, and as 'shift' when the meanings in Dolgan and Sakha do not overlap (anymore). At this point of the analysis, the terminology reflects a synchronic state and no claims are made with respect to the direction of change. Therefore 'broader' could mean semantic extension of a lexical item in Dolgan, but could equally well be due to semantic specification or narrowing in Sakha. It simply indicates that the synchronic meaning in Dolgan is broader than that in Sakha.

CHANGE IN FORM*:* if for a concept Dolgan and Sakha use lexical items with a clearly identifiable common root, but with a difference in phonetic or morphological shape. Differences in phonology are not included, since the phonological systems of Dolgan and Sakha are identical (but see Section 9.3.1.1 for a discussion of differences in the distribution of allophones).

DESCRIPTIVE: if for a concept Dolgan uses a descriptive phrase where Sakha uses a single lexical item, or if Dolgan uses a descriptive phrase different from the one used in Sakha. It is called a calque when the structure of the descriptive phrase is a based on a clearly identifiable source language e.g. Russian.

ABSENCE*:* if a lexical item is absent in Dolgan while it is present in Sakha, most probably due to the fact that the concept has no relevance in the Dolgan culture.

?: if there is too little information available about the difference to classify it into any of the above categories. This can be because the words are not recorded in the available dictionaries or because there is uncertainty about the correctness of the word form.

In the current chapter, only REPLACEMENT and SEMANTIC CHANGE will be discussed in detail. Changes in morphological form is analysed more extensively in Chapter 5, since some of them are not restricted to a few lexical items, but they are part of a more fundamental structural difference between Dolgan and Sakha that has come about through the process of regularisation. The categories DESCRIPTIVE and ABSENCE will be only marginally discussed because differences of these types

seem to be the result of cultural and geographical differences rather than of contact between languages. Denotations for cultural items often disappear from the lexicon as a result the adoption of a different lifestyle due to migrations, colonisation, or other events that may lead to changes in culture. Of course, such concepts can still be referred to by descriptive phrases, which may later be lexicalised to varying degrees. These changes I classify as descriptive. Since it is hard to determine the degree to which such descriptive phrases are lexicalised, and thus whether they belong to the lexicon or are impromptu collocations, I chose not to attribute too much significance to their occurrence for the current purpose.

4.4 RESULTS

4.4.1 QUANTITATIVE ANALYSIS

The quantitative analysis of the lexical differences between Dolgan and Sakha is carried out from two perspectives, the first being the onomasiological approach. In the onomasiological approach (or the approach of naming, see Blank (2003)), the idea is to look how meanings are encoded in language. This means that for each of the 1,500 meanings from the Loanword Typology list the linguistic encoding a comparison in Sakha and in Dolgan is compared. If the encoding is identical across the two languages, there is no reason to conduct further analysis. However, if there is any dissimilarity, the meaning is included in the list of differences, regardless of whether the difference is a replacement, a semantic change, a change in form, a descriptive phrase, or that it is absent. Since the 1,500 meaning list is taken as a point of departure, each meaning can be counted as a difference only once, and so this total number serves as a basis for the onomasiological analysis, as is reflected in Table 4.2.

This is different for the second perspective, that of type of difference. In this approach the aim is to see how the different types of difference (replacement, semantic change, change in form, absence) are distributed over the semantic fields, in particular for the fields of 'the body' and 'kinship'. Since consultants sometimes gave more than one lexical form for a particular meaning (i.e. near synonyms) more than one type of difference can be associated with a particular meaning. Since in this part of the analysis I am interested in the frequency of types of differences in semantic domains, all responses were then taken into account.

This is important to keep in mind, since it explains why there is a higher number of differences (776 to be precise) in this approach, than from the onomasiological perspective (602). To illustrate this with an example, the meaning GRANDDAUGHTER is expressed by the form *sien* in Sakha. In Dolgan *sien* is not used at all and instead this meaning can be expressed by *oγo-m kïːh-a* [child-POSS.1SG girl-POSS.3SG], with the literal meaning 'daughter of my child', or with the Russian word *vnučka*. Both possibilities are considered equally common among Dolgan speakers and are therefore included in the list. Thus, from an onomasiological perspective, this case provides one difference, namely for the meaning of granddaughter. However, from the perspective of type of difference, this example counts as two differences, the one being of the type descriptive, and the other Russian copy.

4.4.1.1 ONOMASIOLOGICAL PERSPECTIVE

Table 4.2 shows the differences in the encoding of the 1,500 meanings across Dolgan and Sakha per semantic field. It shows the total number of meanings included the semantic field, the number of meanings that is encoded differently in the two languages, and the relative percentage of these differences with respect to the total number of meanings in the semantic field. For example, in the semantic field 'emotions and values' 27 out of 48 meanings (56.3%) show a difference in encoding when Dolgan and Sakha are compared, and its position at the top of the table suggests that this field shows most internal diversity. The semantic fields are ranked according to the percentage of differences in descending order.

Table 4.2: Number and percentage of differences in encoding of 1,500 meanings between Dolgan and Sakha

Semantic field	No. of meanings in sem. field	No. of different encodings	% of total no. of meanings in semantic field
Emotions and values	48	27	56.3
Social and political relations	36	20	55.6
Law	26	14	53.8
The house	49	25	51
The body	158	78	49.4
Clothing and grooming	60	28	46.7

Warfare and hunting	41	19	46.3
Agriculture and vegetation	74	34	45.9
Animals	116	53	45.7
Sense perception	49	22	44.9
Kinship	85	37	43.5
Basic actions and technology	79	34	43
Religion and belief	26	10	38.5
The physical world	76	30	39.5
Motion	82	29	35.4
Possession	46	16	34.8
Spatial relations	76	24	31.5
Cognition	62	20	32.3
Time	57	19	33.3
Food and drink	97	32	33
Speech and language	42	13	31
Quantity	40	8	20
Modern world	57	9	15.8
Miscellaneous function words	18	1	5.6
Total meanings	1500	602	40.1

As can be seen from this overview, 602 (40.1%) of a total of 1,500 meanings are expressed differently in Dolgan and Sakha. This seems an unexpectedly high number, considering the fact that the languages are so closely related to each other and have often been described as dialects of the same language. However, this percentage includes all types of difference mentioned above, including phonetic differences. A detailed discussion of the types of differences is presented in Section 4.4.1.2.

The five semantic fields with the highest percentage of differences are ‘emotions and values’ (56.3%), ‘social and political relations’ (55.6%), ‘law’ (53.8%), ‘the house’ (51%) and ‘the body’ (49.4%). Although a direct parallel with the results from the Loanword Typology project cannot be drawn, it is worth drawing attention to the relatively high ranking of ‘the body’ and ‘kinship’. In the Loanword Typology project, ‘the body’ is ranked third lowest when it comes to the proportion of non-native lexical items in this field, with an average of 14.2% cross-linguistically. In addition, body parts occupy a quarter of the Leipzig-Jakarta list of terms that are most resistant to being transferred. Even though these results are about full copies only and not about lexical differences in general, they indicate that cross-linguistically ‘the body’ is a stable semantic field, where a high degree of

change is not to be expected, and even less so for languages as closely related as Dolgan and Sakha. The same cross-linguistic conservatism holds for kinship terminology. Although this semantic field has not made it into the Leipzig-Jakarta list, except for the concept of CHILD, this semantic field is cross-linguistically relatively resistant to the adoption of non-native lexical items, with an average of only 15% of foreign copies (Haspelmath and Tadmor 2009: 64). Against this background, the 43.5% of differences between Dolgan and Sakha in this semantic domain clearly stand out and require more in-depth investigation. These two semantic fields are discussed in detail in Sections 4.4.2.1.1 and 4.4.2.1.2.

4.4.1.2 TYPES OF DIFFERENCE

Table 4.3 summarises the types of difference introduced in Table 4.1 and their frequency of occurrence in the comparison of Dolgan and Sakha. They are grouped by their overarching categories (first column) and are further specified in subcategories (second column). The third column lists for each subcategory its number of occurrences in the data set, and the last column does the same for the number and percentage of the overarching categories. The categories are listed in decreasing order.

Table 4.3: Types of difference: frequency of occurrence

Category	Type of difference	No. of instances	Total for category No.	%
SEMANTIC CHANGE	Broader	332		
	Narrower	14	350	45.1%
	Shift	4		
REPLACEMENT	Russian copy	79		
	Evenki copy	29	129	16.6%
	Unknown	21		
FORM	Morphological change	41	121	15.6%
	Phonetic change	80		
DESCRIPTION	Descriptive	94	95	12.3%
	Calque	1		
ABSENCE	Absence	69	69	8.9%
?	?	11	11	1.5%
Total		775	775	100%

The most obvious observation from Table 4.3 is that the overwhelming majority of differences belong to the category of semantic change. The types of difference 'broader', 'narrower' and 'shift' together comprise almost half of all the total number of lexical differences between Dolgan and Sakha (45.1%). Within the category of semantic change, the type 'broader' accounts for 94.8% of the instances and is thus the most frequent type of difference, not only within this category, but also within all lexical differences in general (42.8% of all differences). This tendency is not restricted to just one or two semantic fields but seems to be pervasive throughout the entire lexicon. In nineteen of the twenty-four semantic fields, lexical differences between Dolgan and Sakha are dominated by semantic change of the type 'broader', and often the occurrence of the second most frequent type of difference is considerably less. The distribution of this type of difference across all semantic fields is presented in Table 4.4. The numbers represent percentages of the total number of differences within the semantic field.

Table 4.4: Types of difference: distribution of subtype 'broader' over semantic fields

Semantic field	% of total no. of differences per semantic field
Emotions and values	72.2%
Possession	65%
Quantity	60%
Social and political relations	58.3%
Speech and language	57.1%
Time	53.8%
Cognition	53.5%
Basic actions and technology	53.4%
Motion	52.9%
Kinship	51.6%
Spatial relations	51.5%
Physical world	48.6%
Religion and belief	46.2%
Sense perception	46.1%
Food and drink	40.5%
Animals	37.5%
Clothing and grooming	37.5%
Warfare and hunting	31.8%
Body	30.7%

House	22.2%
Agriculture and vegetation	20%
Law	18.8%
Modern world	10%
Miscellaneous and function words	0

Representing 16.6% (129 instances) of the differences between Dolgan and Sakha, replacement is the second most frequent category. Table 4.5 shows that within this category, 61.2% of the replacements in Dolgan are copies from Russian, including cases where Russian copies replace a Sakha word, as well as cases where they replace a different Russian word; 22.5% are copies from Evenki; and 16.3% are replacements of unclear origin. In relation to the total number of differences these percentages correspond to 10.2% Russian copies, 3.7% copies from Evenki, and 2.7% of unclear origin. For a detailed discussion of the various cases of replacement see Section 4.4.2.2.

Table 4.5: Types of replacement in Dolgan

Category	Type of difference	% of replacements	% of total no. of differences
REPLACEMENT	Russian copy	61.2%	10.2%
	Evenki copy	22.5%	3.7%
	Unknown	16.3%	2.7%
Total		100%	16.6%

Change in form accounts for 15.6% of the differences, 33.9% of which are differences in morphology and 66.1% differences in the phonological realisation of a lexical form. In 12.3% of the differences in form, Dolgan uses a descriptive strategy where Sakha has a single lexical item. This may be due to contact with Evenki, if it is a calque, i.e. in cases where Evenki uses the same descriptive collocation, or it may be part of ongoing language attrition in Dolgan, during which specific lexical items are lost. The 8.9% coded as absence are primarily concepts concerning agriculture, geographical features, and animals that are not present in the environment of the Dolgan people. These concepts have no relevance for their way of subsistence and hence are not lexically expressed in the language. However, absences are observed in other semantic fields such as body parts, which have nothing to do with a difference in culture or geography. In these semantic domains, the absence of lexical items may have to do with a seeming tendency for generalisation in Dolgan, where terms with less specific meanings

take the place of more specific lexical items of Sakha. This impression is strengthened by the large number of differences of the type 'broader' and the descriptive strategies for concepts for which Sakha uses a single lexical item. While this impression is only based on the restricted set of concepts of the Loanword Typology list, these data evoke an impression of a tendency in Dolgan towards less specific use of lexical items when compared to Sakha.

4.4.2 QUALITATIVE ANALYSIS

4.4.2.1 SEMANTIC CHANGE

Section 4.4.1.2 showed that the majority of lexical differences between Dolgan and Sakha are semantic changes, and that they were classified more specifically as cases where the lexical item in Dolgan covers a wider conceptual space than the same lexical item does in Sakha. Since the quantitative analysis in Section 4.4.1 showed that the high percentage of differences in the semantic fields of 'the body' and 'kinship' is cross-linguistically unusual, these fields deserve to be explored more carefully.

4.4.2.1.1 THE BODY

The semantic field of 'the body' contains 158 concepts, 101 of which are expressed differently in Dolgan and Sakha. A closer look shows that 33 (32%) are semantic change. The complete overview of all types of difference within this domain can be seen in the Table 4.6 below. From this relatively high number of semantic changes, only three instances could possibly be attributed to contact with Evenki. These are the lexical items for the concepts BEAK, FOOT and BRAIN. With respect to the remaining differences, no evidence for language-external motivation could be found, and they are better explained in terms of language-internal semantic variation or change (see Table 4.10 for an example).

Table 4.6: Distribution of changes in semantic field 'the body'

Category	Type of change	No.
SEMANTIC CHANGE	Broader	31
	Narrower	1
	Shift	1
REPLACEMENT	Evenki loan	5
	Russian loan	4
	Unknown	7
DESCRIPTIVE		19
CHANGE IN FORM	Morphological change	6
	Phonetic change	18
ABSENCE		7
?		2

In the following examples, which are presented in tables for reasons of clarity, the first column contains the concepts for which semantic change occurred, the next five columns represent the (proto-) languages for which concepts are compared: Proto-Turkic, to show the semantic pattern that is typical for Turkic languages, then Sakha and Dolgan, which are both Turkic languages but map their lexical items differently onto conceptual space. The two rightmost columns show the lexical items and their semantic distribution for Evenki and where possible for Proto-Tungusic to ascertain if Evenki follows the Tungusic semantic pattern. The different shades of grey indicate the shared cognate forms across languages, and the bold borders group together the languages that map lexical items onto conceptual space in a similar way.

Table 4.7: NOSE/BEAK

Concept	Proto-Turkic	Sakha	Dolgan	Evenki	Proto-Tungusic
BEAK	**tum-š-uk*	*tumus*	*munnu*	*oŋokto*	**hoŋa+kta*
NOSE	**burun*	*murun*	*munnu*	*oŋokto*	**hoŋa+kta*

Table 4.7 shows that in Sakha the concepts BEAK and NOSE are represented by different lexical items. The same pattern can be reconstructed for Proto-Turkic. In contrast, Evenki and the other Tungusic languages, represented by the Proto-Tungusic reconstruction **hoŋa+kta,* use a single lexical item to express both

concepts. Dolgan shows affiliation to both sides: the lexical form *munnu* is related to Sakha *murun*, but its semantic structure matches the Tungusic model, covering the concepts of NOSE and BEAK. The difference in form between *murun* in Sakha and *munnu* in Dolgan is an interesting topic in itself and will be discussed in Chapter 5.

A similar pattern occurs for the concepts LEG, FOOT and SOLE, as illustrated in Table 4.8.

Table 4.8: LEG/FOOT/SOLE

Concept	Proto-Turkic	Sakha	Dolgan	Evenki	Proto-Tungusic
LEG	**aδak*	*ataχ*	*atak*	*halgan*	**palgan*
FOOT	**aδak*	*ataχ/ulluŋaχ*	*ulluŋ*	*hagdïki:*	**hagdï (-ki)*
SOLE	**ultuŋ*	*ulluŋ*	*ulluŋ*	*hagdïki:*	**hagdï (-ki)*

As can be seen from this table, Dolgan unmistakably uses Turkic lexical items, but the semantic distribution of these items lines up with the semantic patterns in Tungusic languages. As in the previous example, the Turkic and Tungusic languages show different groupings of the concepts LEG, FOOT and SOLE. In Turkic languages a single lexical item reflecting **aδak* is used for LEG and FOOT, setting these apart from SOLE. In contrast, Proto-Tungusic and Evenki group together the concepts FOOT and SOLE with a lexical item cognate with **hagdï (-ki)*, and set it apart from LEG. Dolgan employs the Turkic lexical items *atak* and *ulluŋ*, but their semantic distribution corresponds to that of the Tungusic languages, grouping FOOT and SOLE together as opposed to LEG.

The final example is an instance of semantic shift, where the meaning of the Turkic lexical item for BRAIN in Dolgan has shifted its meaning to denote HEAD.

Table 4.9: BRAIN/HEAD

Concept	Proto-Turkic	Sakha	Dolgan	Evenki	Proto-Tungusic
BRAIN	**bejŋ*	*meji:*	*irge*	*irge*	**irgä*
HEAD	**töpe*	*töbö*	*meni:*	*dïl*	**dïlï*

Table 4.9 shows that with respect to the encoding of the concepts BRAIN and HEAD, Dolgan manifests two changes. On the one hand the example displays a

change in semantics of the lexical item *meni:*, which has undergone a semantic shift in Dolgan from denoting BRAIN to meaning HEAD, when compared to Sakha and Proto-Turkic; on the other hand it shows a change in the encoding of the concept BRAIN, resulting in a replacement of the Turkic *meni:* by the Tungusic word *irge* (see also Artemyev 2001: 8). While the replacement itself is palpable, it is impossible to make claims about possible push or pull effects in the sequence of change without historical linguistic data on Dolgan. Was *irge* first copied into Dolgan, leading to the semantic shift of *meni:*? Or had *meni:* in Dolgan first become polysemous for BRAIN and HEAD due to language-internal factors, facilitating the copying of a specific lexical item for BRAIN? Since such historical data for Dolgan are not available, the exact cause of this semantic shift cannot be determined, but from the present use of the Evenki word *irge* in Dolgan it is clear that contact with Evenki played a role in this semantic shift.

For other broader uses of lexical items in Dolgan an explanation in terms of language-internal development is more likely, as is illustrated by the following example featuring the extension of LIVER to BELLY.

Table 4.10: MIDDLE/BELLY/LIVER

Concept	Proto-Turkic	Sakha	Dolgan	Evenki	Proto-Tungusic
INSIDE	*? iš*	*is*	*is*	*dō*	** dō*
BELLY	**karïm*	*is*	*bïar*	*hukite*	**?*
LIVER	**bagïr*	*bïar*	*bïar*	*hakin*	**pa: -kun*

Sakha has one lexical item for INSIDE and BELLY (*is*) and a different word form for LIVER (*bïar*). Dolgan uses the same lexical items, but groups them in a different way; *is* is used for INSIDE only, while *bïar* has the meaning of LIVER and BELLY. However, comparison with both Turkic and Tungusic provides no evidence that this difference must be attributed to contact with Tungusic languages. Evenki has three different lexical items for INSIDE, BELLY and LIVER, so if Dolgan did extend the meaning of *bïar* from LIVER to BELLY, it did not happen according to the Evenki pattern. Similarly, if Sakha extended the meaning of *is* from INSIDE to BELLY, it most likely reflects language-internal change, or potentially contact with a language that was not included into this comparison.

It is remarkable that both Turkic and Tungusic seem to have had specific words for the three concepts and that both Dolgan and Sakha have modified this

pattern, though in different directions. However, cross-linguistically both directions of change are not unusual. According to Wilkins, it is a natural tendency for the meaning of a body part term to "shift to refer to the visible whole of which it is a part" (1981, cited in Wilkins 1996: 273), a pathway that would apply to LIVER → BELLY in Dolgan. The Sakha extension INSIDE → BELLY could be conceived of as an instance of metaphoric change, which is also a common mechanism of semantic change (Geeraerts 2010).

To summarise, out of the 31 semantic extensions in the field of 'the body' there are 3 instances that can be plausibly explained in terms of contact between Dolgan and Evenki. However, while such an analysis is very likely after scrutinising the genealogical and geographical patterns through detailed cross-linguistic comparison, the possibility of internally motivated change cannot be ruled out either, especially since the concepts under consideration are semantically closely related. In addition, the examples discussed above are all instances of semantic extension, which is a common language-internal mechanism of semantic change in viable languages, but even more so in situations of language attrition when the use of specific lexical items gradually diminishes and eventually may be lost. When their functions are taken over by the remaining lexical items, these lexical forms become more polysemous and extend their semantic scope. To put it in perspective, a language-internally motivated semantic extension from one part of the body to a spatially contiguous part, as we see in SOLE → FOOT, is cross-linguistically common (Wilkins 1996: 273), whereas the extension NOSE → BEAK is not. On the basis of a cross-linguistic sample, Wilkins (1981, cited in Wilkins 1996: 273-274) formulated five natural tendencies for language-internal semantic change, number four being:

> It is a natural tendency for an animal part to shift to refer to a person part (e.g. 'snout' → 'nose'; 'beak' → 'face'). (Wilkins 1996: 274)

The fact that in Dolgan we find the opposite direction of change, which is less likely to occur cross-linguistically, in combination with the fact that the new semantic pattern in Dolgan matches the pattern of Evenki, makes contact between these two languages a very plausible explanation of this change.

4.4.2.1.2 KINSHIP TERMS

Within the semantic field of 'kinship', 37 out of 85 (43.5%) concepts are encoded differently in Dolgan and Sakha. In this section I investigate semantic changes in this field, with a particular focus on the concept clusters BROTHER/SISTER, UNCLE/AUNT, MOTHER-IN-LAW/FATHER-IN-LAW, and MAN/HUSBAND-WOMAN/WIFE-FAMILY. For the study of this semantic field, it was not very informative to include Proto-Turkic and Proto-Tungusic reconstructions in the examples, because it turns out that within this time depth the meaning of the proto-forms was often too different from the current meaning to be of help for the reconstruction of a Proto-Turkic or Proto-Tungusic kinship system. However, the terms are referred to in cases where such a reconstruction was possible.

BROTHER/SISTER

Table 4.11 displays the lexical items and their mapping onto conceptual space for the concepts of BROTHER and SISTER in Sakha (and Turkic), Dolgan and Evenki (and Tungusic).

Table 4.11: BROTHER *and* SISTER *in Sakha, Dolgan and Evenki*

	Turkic	Sakha	Dolgan	Evenki	Tungusic
OLDER BROTHER OF ♂	**bi:*	*bi:*	*ubaj*	*aki:n*	**ak'i*
OLDER BROTHER OF ♀	**abaj (?)*	*ubaj*	*ubaj*	*aki:n*	**ak'i*
OLDER SISTER OF ♂		*edʒij*	*edʒij*	*eki:n*	**äkä-i*
OLDER SISTER OF ♀		*aγas*	*edʒij*	*eki:n*	**äkä-i*
YOUNGER BROTHER OF ♂	**ini*	*ini*	*balïs*	*neku:n*	**näkön*
YOUNGER BROTHER OF ♀	**jügürči*	*surus*	*balïs*	*neku:n*	**näkön*
YOUNGER SISTER OF ♂	**baldïz*	*balïs*	*balïs*	*neku:n*	**näkön*
YOUNGER SISTER OF ♀	**badlïz*	*balïs*	*balïs*	*neku:n*	**näkön*

Sakha has an elaborate set of terms to refer to siblings with a three-way distinction, depending on: 1) age relative to ego; 2) gender of the sibling; 3) gender of ego, except in the case of YOUNGER SISTER. That is to say, there are different terms for siblings older or younger than ego, for male or female siblings, and for the sibling of a girl or the sibling of a boy, except when the sibling is a younger sister, which in both cases is *balïs*. This results in the seven-way distinction for siblings as displayed in Table 4.11. This system matches the general Turkic pattern,

and most of the terms are of Turkic origin as well. Only the terms for 'older sister' could be traced back to foreign provenance, of which only *edʒij* with certainty. This term is clearly cognate with the Mongolian form *edʒi*, whereas *aγas* shows, according to Kałużyński, similarity with both the Old Turkic form *äkä* as well as with Mongolic *egeči* (Kałużyński [1968] 1995: 203). Therefore we cannot unambiguously determine whether this term is cognate with Old Turkic term or with the Mongolic term, or that perhaps knowledge of the Mongolic form influenced the shape of the inherited form *äkä*.

In contrast, Dolgan, makes a terminological distinction for which only a subset of the Turkic criteria is relevant, namely 1) age relative to ego, and 2) gender of the older sibling. Unlike Sakha, the gender of the younger sibling and the gender of ego do not play a role in the terminological distinctions. Table 4.11 clearly shows that Dolgan employs Sakha (and Turkic) terms, but that semantic space is divided up in a different way. More specifically, it shows that the semantic organisation of sibling terms in Dolgan exactly matches that of Evenki, and Tungusic more generally: as in Dolgan, Evenki shows a distinction between older and younger siblings and between male and female siblings that are older than ego. Gender of ego does not play a role, nor does the gender of the younger sibling. This comparison of the three languages (and the families more widely) strongly suggests that the Dolgans use Sakha terms, but have restructured the semantic distribution of these terms on the model of the kinship system of the Evenks.

UNCLE/AUNT

Since the labeling of the kinship relations of AUNT and UNCLE in Sakha and Dolgan is somewhat more complicated than for the previous example, the distribution of terms is given for Sakha, Dolgan and Evenki separately (see Tables 4.12, 4.13, 4.14), building up towards the complete picture. An overview of the entire system, in which the similarities and differences are clearly seen, is given in Table 4.15. In Sakha, aunts and uncles are categorised depending on the gender of ego's parent (see Table 4.12): father's brother is called *abaγa* and mother's brother is called *ta:j*. The relative age of the uncle/aunt to ego's parent has no influence on the choice of terminology. Although it is not possible to make any claims about a general Turkic system with certainty, it seems to have been similar to the one in Sakha.

Table 4.12: AUNT/UNCLE *in Sakha*

	Age relative to parent	
	older	younger
BROTHER OF FATHER	*abaγa*	*abaγa*
BROTHER OF MOTHER	*ta:j*	*ta:j*
SISTER OF FATHER	*edʒij*	*edʒij*
SISTER OF MOTHER	*ta:j edʒij*	*ta:j edʒij*

According to Tenishev (2001: 296) *ta:j* can be traced back to **taγa*, meaning 'uncle from mother's side' (i.e. brother of mother), not mentioning relative age to ego's parent as a relevant criterion and thus matching Sakha. With respect to *abaγa* opnions differ as to whether the term has Turkic (Tenishev 2006: 228) or Mongolic origin (Kałużyński [1962] 1995: 54), but in either case the meaning is 'uncle from father's side' (i.e. brother of father), again not mentioning relative age as a distinguishing feature. The comparison of the category of AUNT between Sakha and other Turkic languages is less straightforward, since for SISTER OF FATHER the Mongolic term *edʒij* is used (see previous example), and for SISTER OF MOTHER a combination of a Turkic and Mongolic term. However, despite this deviation from Turkic in the lexical *forms*, the kinship *categories* in Sakha could still match the Turkic pattern, if we assume a symmetrical relation between identification of uncles and aunts. However, this cannot be determined with certainty.

In Dolgan, on the other hand, the gender of ego's parent is irrelevant but instead the age of the parent's sibling relative to ego's parent is the deciding factor.

Table 4.13: AUNT/UNCLE *in Dolgan*

	Age relative to parent	
	older	younger
BROTHER OF FATHER	*ehe*	*uba/ubaj*
BROTHER OF MOTHER	*ehe*	*uba/ubaj*
SISTER OF FATHER	*ebe*	*edʒij*
SISTER OF MOTHER	*ebe*	*edʒij*

The parent's brother older than ego's parent is called *ehe* regardless of whether that is on the father's or mother's side. Similarly, parent's brothers younger than ego's parent (either mother or father) are called *uba/ubaj*. The same pattern

applies for parent's sisters, labeled *ebe* and *edʒij*. An identical pattern is found in Evenki, as is shown in the table below.

Table 4.14: AUNT/UNCLE *in Evenki*

	Age relative to parent	
	older	younger
BROTHER OF FATHER	*ama:ka*	*aka/aki:n*
BROTHER OF MOTHER	*ama:ka*	*aka/aki:n*
SISTER OF FATHER	*ene:ke*	*eki:n*
SISTER OF MOTHER	*ene:ke*	*eki:n*

The lexical forms of Evenki are unrelated to the ones used in Dolgan, but their distribution in conceptual space is exactly the same. Ego's uncles and aunts are labeled differently depending on their relative age to ego's parent, as is the case in Dolgan. Comparison with other Tungusic languages shows that this pattern is common in the entire family: Proto-Tungusic **ama:ka:n* referred to '(mother's) older brother' (Doerfer 2004: 68), and although for *ańaka* no reconstruction is given, the fact that related forms occur in 13 Tungusic languages and dialects is good evidence that this term, and its meaning of '(father's) elder sister' are widespread across the Tungusic family. Now combining the patterns in Sakha, Dolgan and Evenki in one table, we see the following picture.

Table 4.15: UNCLE/AUNT in Sakha, Dolgan and Evenki

	Sakha	Dolgan	Evenki
OLDER BROTHER OF FATHER	*abaγa*	*ehe*	*ama:ka*
OLDER BROTHER OF MOTHER	*ta:j*	*ehe*	*ama:ka*
OLDER SISTER OF FATHER	*edʒij*	*ebe*	*ene:ke*
OLDER SISTER OF MOTHER	*ta:j edʒij*	*ebe*	*ene:ke*
YOUNGER BROTHER OF FATHER	*abaγa*	*uba/ubaj*	*aka/aki:n*
YOUNGER BROTHER OF MOTHER	*ta:j*	*uba/ubaj*	*aka/aki:n*
YOUNGER SISTER OF FATHER	*edʒij*	*edʒij*	*eki:n*
YOUNGER SISTER OF MOTHER	*ta:j edʒij*	*edʒij*	*eki:n*

While the identical distribution of terms in Dolgan and Evenki itself is suggestive of contact influence, this idea becomes even more appealing as the semantic details of the Dolgan terms are put under the magnifying glass. Table 4.15 shows

that of all the terms for AUNT and UNCLE in Dolgan, only *edʒij* is shared with Sakha, be it with only a partially overlapping denotation. In Dolgan, *edʒij* denotes a younger sister of ego's parent (regardless of whether father or mother), whereas in Sakha it is the term for the sister of ego's father (regardless of whether older or younger), reflecting once again the importance of relative age to ego's parent in Dolgan versus the importance of gender in Sakha.

While it is not visible in the table above, the other terms used in Dolgan are also shared with Sakha. In Sakha and in Dolgan, *ehe* is used to refer to 'grandfather', *ebe* for 'grandmother' and *ubaj* for 'older brother' (see Table 4.11). However, in Dolgan the semantic coverage of these lexical items is broader than in Sakha, including the meanings of uncle and aunt as well and importantly, the same polysemy is found in Evenki.

Table 4.16 Polysemy of terms used for UNCLE and AUNT in Dolgan and Evenki

	Dolgan	Evenki	Gloss
OLDER BROTHER OF FATHER	*ehe*	*ama:ka*	'grandfather'
OLDER BROTHER OF MOTHER	*ehe*	*ama:ka*	
OLDER SISTER OF FATHER	*ebe*	*ene:ke*	'grandmother'
OLDER SISTER OF MOTHER	*ebe*	*ene:ke*	
YOUNGER BROTHER OF FATHER	*uba/ubaj*	*aka/aki:n*	'older brother'
YOUNGER BROTHER OF MOTHER	*uba/ubaj*	*aka/aki:n*	
YOUNGER SISTER OF FATHER	*edʒij*	*eki:n*	'older sister'
YOUNGER SISTER OF MOTHER	*edʒij*	*eki:n*	

In Evenki, the word used for older brother of ego's parent, *ama:ka*, is the same as the word for grandfather, and a younger brother of ego's parent, *aka/aki:n*, also means older brother of ego. The same holds for the terms for sisters of ego's parents: *ene:ke* means older sister of ego's parent and grandmother, while *eki:n* means younger sister of ego's parent and older sister of ego. In Dolgan the pattern is identical. *Ehe* is older brother of ego's parent but also grandfather, and *ebe* is older sister of ego's parent but also grandmother. *Uba/ubaj* is younger brother of ego's parent and older brother of ego, and *edʒij* is younger sister of ego's parent and older sister of ego. Thus, the organisation of referential terms for aunt and uncle, as well as the semantic details of the terms chosen for this purpose strongly suggest that the similarities between Dolgan and Evenki are no coincidence but that they have developed as a result of contact between the two populations.

MOTHER-IN-LAW/FATHER-IN-LAW

Table 4.17 displays the terms used to refer to parents-in-law. In Sakha, the terms for parents-in-law are organised both according to the gender of ego and the gender of the parent-in-law. A male ego refers to his parents-in-law as *ayа kїlїn* and *ińe kїlїn*, a female ego uses the terms *tojon* and *χotun*. While the terms used in Sakha may all be of Turkic origin[2], it is not clear whether this particular system of reference to parents-in-law is typical for the Turkic language family, since all terms originally had a rather different meaning. For example *kїlїn* < **qayin* 'wive's relatives' (Tenishev 2001: 309), *tojon* < *tojїn* 'monk' (Pekarskij [1907 - 1930] 1958-1959: 2706) and *χotun* < **qatyn* 'wife' (Tenishev 2001: 296).

Table 4.17: FATHER-IN-LAW/MOTHER-IN-LAW in Sakha, Dolgan and Evenki

	Sakha	Dolgan	Evenki
FATHER-IN-LAW OF ♂	*ayа kїlїn*	*kїnnї*	*etki:*
FATHER-IN-LAW OF ♀	*tojon*	*kїnnї*	*etki:*
MOTHER IN LAW OF ♂	*ińe kїlїn*	*ińe kїnnї*	*atki:*
MOTHER IN LAW OF ♀	*χotun*	*ińe kїnnї*	*atki:*

In Dolgan on the other hand the gender of ego does not play a role. A male and a female ego both use the same terms to refer to their mother- and their father-in-law. The differences in the choice of terms depend solely on the gender of the parent-in-law. As in the previous examples, Evenki uses unrelated lexical items, but their semantic distribution is the same as in Dolgan. In addition, comparative Tungusic etymology shows that this system of reference is deeply rooted in the family: *etki:* < *ekk'in* 'father in law' and *atki:* < *atk'i* 'mother in law' (Doerfer 2004: 100, 295).

Although linguistic data alone is not sufficient to postulate conclusions with respect to admixture patterns between Dolgans and Evenks in the past, they are an important component within the broader picture including historical, anthropological and genetic evidence. A case in point in this context is the semantic extension of *kїlїn*, which means 'parent-in-law of male ego' in Sakha, and for which the preceding *ayа* 'father' or *ińe* 'mother' specifying the gender of the

[2] For *tojon* and *qatun* also other origins are suggested.

parent. In Dolgan *kïľïn* is extended (after a morphological change leading to the form *kïnnï*, see Chapter 5 for details) to denote 'parent-in-law' regardless of the gender of ego, but with a specification of *ińe* 'mother' for 'mother-in-law'. The fact that 'parent-in-law' from the male perspective has been kept and extended in Dolgan could arguably be indicative of a pattern of Evenki women marrying into the Dolgan community. This would be compatible with the tradition of patrilocality in both Dolgan and Evenk communities (Ventsel 2005: 152, personal observation), as well as with the percentage of sharing of mtDNA haplotypes (see Sections 2.6.2, 2.6.4). If an Evenki woman married a Dolgan man and presumably began to learn the Dolgan language, she would have heard most Dolgan speech within her new Dolgan family and from her husband. Given the fact that people normally speak from their own perspective, this means that she would have heard *kïľïn* (parents-in-law from the male perspective) more frequently than *χotun* and *tojon* (the Sakha terms for parents-in-law from a female perspective, which may have been used before contact with the Evenks). The husband would have used *aγa kïľïn* and *ińe kïľïn* to refer to her parents (i.e. his parents-in-law), while for *her* parents-in-law (i.e. his own parents) the husband would have used *kergenner* 'parents'. On hearing *kïľïn* being used by her husband for parents-in-law, the Evenki woman, as a second language learner of Dolgan, may have identified this term with the Evenki terms *etki:* and *atki:*. Through interlingual identification she may have projected the semantic properties of the Evenki terms onto the Turkic word, leading to a generalisation of *kïľïn* to denote 'parent-in-law' from the male as well as female point of view.

HUSBAND/WIFE/WOMAN/MARRY

The final example concerns the conceptually related terms for husband, wife and marriage. For this set of concepts Evenki influence is not as compelling as in the previous examples but as will be shown, it could nonetheless help explain the difference in semantic reorganisation between Dolgan and Sakha. It is necessary to point out that this semantic area shows a variety of terms to refer to a single concept, all with their own shades of meaning, especially in Sakha. It is therefore difficult to define a single lexical item as *the* word for husband or for wife. For the same reason, a comparison with other Turkic and Tungusic languages proved not helpful for this example. For the purpose of clarity, only those lexical items that are shared by two or more languages are represented in the table below, but for

the purpose of completeness, the alternatives for HUSBAND and WIFE and FAMILY in Sakha are listed in a footnote.

Table 4.18: HUSBAND/WIFE/FAMILY/MARRY in Sakha, Dolgan and Evenki

	Sakha[3]	Dolgan	Evenki
MAN	*er*	*er*	*edï:*
HUSBAND	*kergen*	*er*	*edï:*
WOMAN	*dʒaχtar*	*dʒaχtar*	*asi:*
WIFE	*kergen*	*dʒaχtar*	*asi:*
FAMILY	*kergen*	*kergen*	*kergen*
MARRY	*kergennen*	*erden*	*edi:le:mi:*
		dʒaχtardan	*asi:la:mi:*

To start with the first data cell in Table 4.18, *er* in Sakha is used for MAN, whereas in Dolgan it has the added meaning of HUSBAND. To be fair, this is a possible meaning in Sakha as well, but it is not very common[4]. However, Evenki also has a single lexical item to refer to these two concepts, and this model could have reinforced the 'husband' aspect of the meaning of *er*. This scenario is supported by the fact that a similar situation applies for WOMAN and WIFE. Dolgan, as well as Evenki, use one lexical item to express both concepts, while for Sakha I have no evidence that *dʒaχtar* 'woman' is used with the meaning of 'wife'.

Kergen is pervasive in all three languages, but has a more limited meaning of 'family' in Dolgan and Evenki, as opposed to 'family', 'wife' and 'husband' in Sakha. Originally the word comes from Mongolic *gergen* 'wife, married woman' (Kałużyński [1962] 1995: 156, Lessing 1995: 379a), a meaning that has been kept in Sakha but has over time extended to cover also 'husband' and 'family'. Although not all of the details of this semantic change can be established, the main point here is the observation that this term has the same semantic distribution in Dolgan and Evenki, and that this is different from the semantic distribution in Sakha.

---

[3] *Sakha alternatives:* HUSBAND*: er* - 'man'*, oγonńor* - 'old man'*,* WIFE*: ojoχ* - 'woman' *eme:χsin* - 'old woman', *FAMILY: ïal* - 'family' 'homestead' 'neighbour' *dʒon* - 'people' 'family'.

[4] In fact the most common way to refer to one's husband in Sakha is *oγonńor*, at least in the district of Tattaa where I conducted my fieldwork. However, since this term is irrelevant for the current comparison it has been left out.

The meaning of 'husband' and 'wife' for *kergen* in Sakha has probably facilitated the derivation of the verb 'to marry' as well. The verb root

(1) *kergennen*
kergen-LA: -(I)n
spouse–VBLZR-RFL
'to marry'

could literally be translated 'to spouse oneself' or in other words 'to marry'. In Dolgan and Evenki the word for 'to marry' is derived from 'man'/'husband' or 'woman'/'wife' depending on the gender of the person who marries. According to Pekarski ([1907-1930] 1958-1959), Sakha has these words as well, but clearly modern Sakha people would use *kergen* first, whereas the Dolgans would not.

4.4.2.2 REPLACEMENT

As mentioned in Section 4.3.2, three categories of replacement were identified in Dolgan, including copies from Russian, copies from Evenki and lexical items of unknown origin. Copies from Russian have entered the language at different stages, whereby a rough division can be made between the pre-Soviet and the Soviet period. As may be recalled from Chapter 2, the nature of the relations between Russians and indigenous peoples was different during each of these stages. This is important to keep in mind because it may have had consequences for the types of change we see in the lexicon as well as in other domains of the Dolgan language today. Russian copies are discussed in Section 4.4.2.2.1; copies from Evenki in 4.4.2.2.2 and lexical items of unknown origin are briefly mentioned in 4.4.2.2.3. An overview of the kinds of replacement, their absolute numbers and percentages is repeated in the table below.

Table 4.19: Replacement in Dolgan

Type of difference	No. of instances	% of replacements	% of total no. of differences
Russian copy	79	61.2%	10.2%
Evenki copy	29	22.5%	3.7%
Unknown	21	16.3%	2.7%
Total	129	100%	16.6%

4.4.2.2.1 RUSSIAN COPIES

On the comparability of the Loanword Typology list for Dolgan and Sakha

Russian copies constitute the largest proportion of replacements in Dolgan. The 79 Russian copies referred to in Table 4.19 make up for 61.2% of all replacements, which corresponds to 10.2% of the total number of lexical differences between Dolgan and Sakha. As was specified earlier, these 79 Russian replacements mean that 79 concepts of the Loanword Typology list are expressed by a lexical item of Russian origin in Dolgan, where Sakha a) uses a non-Russian word or b) uses a Russian word that is different from the one used in Dolgan[5]. However, in practice the overwhelming majority turned out to be of the first type. To quantify this statement, in 68 out of 79 cases (86.1%), the Russian copy in Dolgan replaces a Sakha word of non-Russian origin, leaving only 13.9% for the second scenario. To eliminate confusion, it needs to be emphasized that these 79 differences are counted from an onomasiological perspective. This means that they include cases of polysemy, where a single Russian term was used to express more than one concept in the Loanword Typology list. This in turn means that the absolute number of Russian copies is slightly lower.

While these percentages provide information about the distribution of Russian copies within the subset of the Dolgan lexicon covered by the Loanword Typology list, a thorough comparison of the wordlist for the two languages shows that it is impossible to make any claims about differences in the overall proportion

[5] This number does not reflect the total number of Russian lexical items. For the current purpose only the differences between Dolgan and Sahka are counted, and therefore meanings for which both languages use the same copy from Russian were not included in the calculations. Therefore, the total percentage of Russian copies is higher.

of Russian copies between the two languages. The lexicon of Dolgan, as well as of Sakha, contains a large set of Russian words, but as it turns out these sets may not, and often do not, overlap. That is, Sakha may employ Russian copies for a subset of concepts of the Loanword Typology list that only marginally overlaps with the subset for which Dolgan uses Russian words. This result is further complicated by the questionable integration of some Russian items into the native lexicon, in other words, whether they are copies that have been accepted by the entire speech community or whether they are only a nonce-borrowing produced as a sign of willingness to fill the slot in the elicitation task. It is in the nature of elicitation data that the results depend heavily on the language proficiency of the language consultants, on the attitude towards (preservation of) the language, or on the reluctance to accept foreign copies as an integral part of the language. While a simple mechanical count would lead to the conclusion that Dolgan and Sakha have exactly the same overall proportion of Russian copies, namely 16.3%, a closer look reveals that this number has come about for both languages in different ways. From my elicited data for Dolgan it becomes clear that my language consultant put in much effort to keep the lexicon as authentically Dolgan as possible. In contrast to the consultant for Sakha, she preferred to leave entries such as 'beaver' or 'oak tree', which are irrelevant in both languages, empty rather than filling them with Russian words. This desire to fill out every entry in the Loanword Typology list accounts for 31 Russian copies in Sakha, so if they were excluded from the comparison, the percentage in Sakha would be slightly lower than in Dolgan. In a similar vein, my Dolgan informant also avoided Russian words by using Dolgan descriptive expressions, e.g.

(2) *die* *ürdü-te*
house top-POSS.3SG
'roof'

instead of Sakha *kïrï:sa* 'roof', which is based on Russian *krïša*. Also, she would prefer to find archaic words such as *hurbuk* 'wooden peg' instead of Sakha *bi:nte*, from Russian *vint* 'screw', for concepts which in everyday speech would be expressed through Russian words. While this yields valuable information about archaic words and their cultural connotations, it does not give an entirely realistic representation of the contemporary Dolgan lexicon as used in everyday speech, and, more importantly, makes direct mapping to the Sakha word list problematic. Thus the divergent results for Russian copies in Dolgan and Sakha are caused for a

large part by the optional employment of these items, as well as by the differences in language attitude of my Dolgan consultant and the consultant for Sakha.

With this knowledge in mind, we can say that in many cases where Sakha uses a Russian copy and Dolgan does not, the two realisations may be used interchangeably: Dolgan could use the Russian copy, and Sakha could use a descriptive phrase like Dolgan to express the concept. The reverse, however, is not true: for those entries where Dolgan uses a Russian word and Sakha does not, the Sakha word is often unknown to the Dolgans or it is used with a different meaning. It is these words that constitute the list of differences between Dolgan and Sakha that are discussed below.

Distribution of Russian copies in the Dolgan lexicon

Russian replacements (i.e. copies from Russian that have replaced a Sakha word or a different Russian word) are not limited to particular semantic domains but are pervasive throughout a large part of the Dolgan lexicon. However, they are not equally distributed over the 24 semantic fields, and their proportions vary from 18.3% Russian copies in the most affected semantic domains to 0% in the least affected ones. The five semantic fields with the highest proportion of such Russian replacements are 'the house' (18.3%), 'clothing and grooming' (13.3%), 'warfare and hunting' (12.2%), 'agriculture and vegetation' (9.5%) and 'kinship' (9.4%).

For most of these semantic domains their high ranking is not unexpected when compared to Sakha, nor from a cross-linguistic perspective. Although the purpose of the Loanword Typology project was different from the current purpose to quantify the difference in copied lexical items between two languages, it may still be curious to view this specific result against the cross-linguistic picture of 'borrowability' to get an impression of what is typical and atypical in the distribution of foreign copies across semantic fields. Both in Sakha and cross-linguistically 'the house', 'clothing and grooming' and 'agriculture and vegetation' fall within the top five, while 'warfare and hunting' is in positions seven and eight respectively (Tadmor 2009: 64, Pakendorf and Novgorodov 2009: 507). The fact that these semantic fields also appear high on the list of differences between Dolgan and Sakha is an indication that Dolgan took the trend in Sakha a step further: those fields that are generally prone to influence from foreign languages have

experienced even more influence from Russian, either due to more intense contact in the past, or to the increased dominance of Russian in recent times.

The only exception is 'kinship', which shows stronger foreign influence in Dolgan both when compared to Sakha and the cross-linguistic average. As mentioned earlier, 'kinship' in Dolgan displays the fifth highest proportion of (Russian) replacements, which is remarkably higher than the 19th position this semantic field occupies in Sakha or 21st cross-linguistically. Even though this may seem striking, the differences themselves are not all that significant, since the Russian terms are all used alongside native Turkic lexical items.

The five semantic fields in which the lowest proportion of Russian replacements are found are 'motion', 'religion', 'speech and language', 'quantity' and 'miscellaneous and function words'. Four of these fields do not show any Russian copies at all, only 'motion' employs *bolot* for 'raft' from Russian *plot*, instead of the Sakha word *a:l*, which was not known to my informant.

The fact that 'religion' ranks so low in Dolgan, while it is the highest ranked domain in terms of copying cross-linguistically, is explained by the fact that Dolgan does indeed employ copies from Russian in this semantic field, but since they are identical to the Russian copies that are used in Sakha they do not classify as a difference.

Russian copies replacing Sakha words

The Russian copies in Dolgan that replace Sakha words can be divided into three types: 1) the concept and lexical item are both foreign, while in Sakha the concept and the lexical item are both native; 2) the concept and lexical item in Dolgan are both foreign, while in Sakha the concept is foreign, but the meaning of a native lexical item has been extended to express it; 3) the concept is native but the lexical item is foreign, while in Sakha both concept and lexical item are native.

The first type concerns concepts that were known in the traditional lifestyle of the Sakha, but lost their relevance when groups of Sakha began to move north and adopted a different lifestyle. This mainly applies to semantic fields such as 'agriculture', 'animals', 'the house' and 'warfare and hunting'. Assuming that this is what happened, it is not surprising that many words connected to these domains have changed in Dolgan. Concepts that belonged to the 'old' Sakha lifestyle of cattle breeding and life in the taiga lost their relevance and related lexical items

were lost, while concepts related to the 'new' Tungusic lifestyle of reindeer herding in the tundra as well as trade with Russians gained importance and had to be added to the lexicon. Often the lexical items were adopted from the language spoken by the people who introduced the concepts, whether they were Evenks (for the lexicon of reindeer terminology) or Russians (for the lexicon of trade). Thus, Russian copies have been entering the language over a long period of time, starting in the 17th century and continuing today. Interestingly, some of the 'forgotten' concepts, especially in the semantic field of agriculture, regained importance during the Soviet regime, when cultural contact with the Russians was particularly intense. The examples below show that Dolgan employs many Russian terms where Sakha has retained the native lexical items. In some cases, the Russian terms have undergone semantic change, as in *ferma*, which has extended from 'farm' in Russian to 'farm' and 'stable' in Dolgan, and *document*, which has extended from 'document' in Russian to also mean 'driver's licence', see Table 4.20 and 4.21 below.

Table 4.20: Russian terms where Sakha has native term and native concept

Concept	Sakha	Dolgan	Russian
COW	*ïnaχ*	*koruoba*	*korova*
STABLE	*χoton*	*ferma*	*ferma* (farm)
PITCHFORK	*atïrdʒaχ*	*vi:la*	*vila*
BUTTER	*arï:*	*sili:be maslata*	*slivočnoe maslo*

The second type is represented in Table 4.21 and exemplifies cases where the concept is foreign to both Dolgan and Sakha, but where Dolgan has adopted a lexical item from Russian and Sakha employs a native word, the meaning of which has expanded to cover the new concept.

Table 4.21: Russian term where Sakha has extended native term and foreign concept

Concept	Sakha	Dolgan	Russian
PLOUGH	*χorut (dig)*	*pahajda:*	*paxat'*
LOCK	*χata:hïn (bolt)*	*homuok*	*zamok*
DRIVER'S LICENSE	*köŋül (freedom)*	*dokument*	*dokument*
WALL	*erkin (side)*	*istiene*	*stena*

Aside from copies that came with Russian concepts, Russian lexical items do also occur for concepts that are completely independent of Russian contact and that would have been relevant long before that. This situation was summarised as type three above. Since in these cases the practical need for a Russian word is not so obvious, they could be more illuminating with respect to other potential motivations for the adoption of such foreign lexical items. It may reflect aspects of the possible relationship between Russians and Dolgans in the past and will therefore be discussed in more detail below, despite their small number. From a total of 79 Russian copies in Dolgan that are different from Sakha, six fall into this category.

Table 4.22: Russian form for native concepts

Concept	Sakha	Dolgan	Russian
MOSQUITO	*bïrdaχ (kuma:r)*	*kuma:r*	*komar*
ROOT	*silis*	*kakuora*	*kokora* (hook)
SWAMP	*dʒebere, kuta*	*namuluoχ*	*navolok* (washland)
COLOUR	*öŋ*	*hibiet*	*tsvet*
WEST	*arγa:*	*hapad*	*zapad*
DANDRUFF	*χoγoho*	*perxot'*	*perxot'*

The most astonishing result is to find the concept MOSQUITO in this list. Of all living creatures on the Taimyr, these fellows are surely the most prominent and cannot possibly escape the attention of anyone who sets foot on the Peninsula. Importantly, their presence is completely independent of the presence of Russian colonisers. Given the fact that during summer months they dominate the lives of humans and animals it is rather surprising to find that the Dolgan people do not use a native word to refer to this omnipresent phenomenon. Instead of using the Sakha word *bïrdaχ*, they use *kuma:r*, which is clearly cognate with Russian *komar*. According to the dictionary, it is possible to use *kuma:r* in Sakha as well, but during my own visits to the Sakha, *bïrdaχ* was always the default translation and it is used far more frequently.

Silis 'root' exists as *hilis* in Dolgan as well, but according to the Dolgan people I spoke with it has the meaning of 'leaf'. *Kokora* in Russian means 'hook' and can be used in Sakha too to mean 'hook' or 'tree with a hooked root' (Anikin, 2003: 277), while in Dolgan it is taken to mean 'root' in general. So *kakuora* must be classified

as a replacement for *silis* in the meaning of 'root', but with a semantic change through polysemy from 'hook' to 'tree with a hooked root' to 'root' in Dolgan.

A confrontation with swamps is also unavoidable when living on the Taimyr, but instead of using Sakha *dʒebere* (from Mongolic) 'swampy mud', or *kuta* (from Evenki) 'quagmire', the Dolgans refer to this phenomenon with *namuluok* from Russian 'navolok' meaning a low place on the riverbank where the river leaves its sediments after overflowing. According to Anikin (2003: 387), *namïlïaχ* in Sakha means 'the transfer of a boat by portage from one water body to another', so it may occur in Sakha, but with a different meaning.

Since these lexical items (can) occur in both Dolgan and Sakha, it is impossible to tell with certainty whether they were copied once into the common ancestor language of Dolgan and Sakha, or whether they were copied at a later stage into both languages independently. According to the first scenario the Russian words would have been copied into a hypothetical common language D/S (Dolgan/Sakha) before groups of Sakha people began to migrate to the north and their language began to diverge. The differences in use and meaning that we see today could then be due to language-internal changes that occurred in the individual languages after their separation. Alternatively, the difference in meaning could be due to the fact that the Russian lexical items were copied into both languages independently and in a different context, and therefore with different connotations.

This category also includes cases where Sakha does not use a Russian word. *Öŋ* 'colour' occurs in Sakha, as well as in the Dolgan dictionary by Stachowski (1993: 199), but a preceding asterisk indicates that its occurrence in Dolgan is not certain. The language consultants I worked with did not know the word and produced the Russian *hibiet* instead. DANDRUFF in Dolgan also has no indigenous term but is referred to by the Russian word *perxot'*.

Russian copies for concepts of orientation and cardinal direction are interesting since these concepts seem relevant to everyone, and in particular to nomadic people in a polar desert with very little help of landmarks. Instead of having specific lexical items to express direction, Dolgan uses descriptive terms for all directions, whereas Sakha has a full system of Turkic terms. In addition, the term *hapad* 'west' has been copied from Russian *zapad*, which may have to do with the fact that the west is, besides a cardinal direction, also a frequently mentioned socio-political unit.

To summarise, of the three types of Russian copies discussed above, the cases where a Russian copy has been introduced to refer to an already existing concept are most intriguing because there seems to be no objective need for a foreign term. The fact that Russian terminology has pervaded as far into Dolgan as to cover culturally independent concepts such as MOSQUITO or COLOUR could be explained in several ways. Theoretically, it could be a reflection of a numerical dominance of Russian immigrants in the 17th century who introduced these words into the Dolgan-speaking area. The first settlers are said to have interacted with the native population considerably, and to have learned the local language. However, there is no evidence that the number of second language learners of Dolgan was overwhelming. While there were certainly Russian individuals who learned Dolgan as a second language, they are often reported in the literature to have integrated completely with the native population (Dolgikh 1963: 121) and become indistinguishable. While this may be exaggerated, it indicates that they probably knew Dolgan well enough to pick up on words like MOSQUITO or COLOUR. Second, if these Russians spoke Dolgan so well that they practically merged with native speakers, there was no large and distinguishable community of second language speaking Russians, which would be necessary for Russian substrate influence to become established in the Dolgan speech community. Therefore, an explanation in terms of imposition due to numerical dominance of L2 learners of Dolgan is unlikely. More likely the adoption of such words reflects dominance of the Russian language in terms of prestige rather than in number. Maybe the number of Russian people was not large enough to establish substratum effects, but the increasingly dominant status of their language from the 19th century onwards could be a motivation for the introduction of Russian terms in Dolgan, even in the basic lexicon.

Finally, there is the possibility that these copies are relatively recent, and that they are part of the language attrition that is observed as a result of the on-going shift to Russian. However, this scenario is rather unlikely considering the degree of phonetic integration of these Russian words into the phonological system of Dolgan, which is much stronger in old words than in recently copied ones.

Replacement of Russian words

For 12 concepts both Dolgan and Sakha use a Russian word, but they are different. In most cases this seems to have happened randomly, but in some cases it may reflect linguistic conservativeness of Dolgan when compared to Sakha (see Artemyev 2001a: 9), in that Dolgan uses copies of older dialectal Russian terms, whereas Sakha uses words from modern, literary Russian, as is illustrated in Table 4.23.

Table 4.23: Different Russian copies for the same concept in Dolgan and Sakha

Concept	Sakha	modern Russian	Dolgan	dialectal Russian
BEAUTIFUL	*kïrahïabaj*	*krasivij*	*baskuoj*	*boskoj*
CALENDAR	*χalenda:r*	*kalendar'*	*paskal*	*paskal'*

4.4.2.2.2 EVENKI COPIES

General remarks

The number of lexical copies from Evenki is remarkably low given the close relationship between Dolgans and Evenks reported in the literature (see Chapter 2), and the characterisation of Dolgan as 'Turkic grammar with Evenki lexicon' or even as a 'creole' (Ziker 1998: 102). As was shown in Table 4.19, 3.7% of the overall differences between Dolgan and Sakha are copies from Evenki, which corresponds to 22.5% of all replacements. Statements that the Dolgans originated from different Tungus clans (Popov [1931] 2003: 60) or that they are 'Yakuticized Evenks' (Dolgikh (1935) cited in Anderson 2000: 86) suggest a very close connection with the Evenks, which could be expected to have had its repercussions on the language. While the current opinion on the origins of the Dolgans is more nuanced, it is undisputed that the Evenks have played an important part in the formation of the Dolgan people, and that there was substantial contact between the two populations (see Chapter 2 for details).

It is important to remember that if a broader range of semantic fields had been included in the analysis, the outcome might have been different. Culturally specific vocabulary, in particular terminology related to e.g. reindeer herding, hide preparation, sleigh riding would have yielded a higher number of copies from

Evenki since it is a known fact that most reindeer terminology in Dolgan was adopted from Evenki. Adoption of culturally specific terminology can thus be an indication of a change in culture, and since these kinds of copies can enter a language also in scenarios of rather superficial contact (Ross 2003: 193) they are not necessarily helpful in the study of population history and possible admixture of peoples. Foreign copies in non-cultural vocabulary occur less easily (Hock & Joseph 1996: 245) and are therefore a more reliable marker of in the study of contact. So even though the number of Evenki copies is higher in Dolgan than in Sakha, the claim that the proportion of Evenki copies constitutes the main difference between Dolgan and Sakha seems to be, with only 3.7% difference, an exaggeration.

Distribution of Evenki copies in the Dolgan lexicon

Since the overall number of replacements from Evenki is not very high, quantitative results for their distribution across semantic fields do not carry much significance. Nevertheless, I consider it worthwhile to give an impressionistic picture on the basis of the available data, from which it appears that the Evenki copies that replace a Sakha word are more restricted in their distribution across the included semantic fields than copies from Russian. The highest percentages occur in the semantic fields of 'the house' (6.1%), 'kinship' (5.9%), 'animals' (3.4%), 'clothing and grooming' (3.3%), 'the body' (3.2%), 'the physical world' (2.6%). Obviously, Evenki has had no influence in domains having to do with modern developments or modern social organisation, such as law, and social and political relations. These spheres are dominated by Russians and the labels for concepts related to these domains were introduced in Russian. More surprising is the fact that Evenki has not left its traces in the domain of 'food and drink', even though this would fall in line with areas such as 'the house', 'clothing and grooming' and 'the body'.

In more than one third of the Evenki replacements (11 out of 29), the Evenki word and the Sakha word are mutually exclusive: the Sakha word does not exist in Dolgan and the Evenki word is not found in Sakha. As can be seen in Table 4.24, for two of the examples this exclusivity exists because Sakha did not seem to have a lexical item for the concept in question.

Table 4.24: Copies from Evenki, Sakha word does not exist

Concept	Sakha	Dolgan	Evenki
SPIDER	*oγuj*	-	-
	-	*ata:ki*	*ataki:*
BUTTERFLY	*ürümečči*	-	-
	-	*lörüö*	*le:re:*
ELK	*tajaχ*	-	-
	-	*tuoki:*	*to:ki:*
YARD	*olbuor/telgehe*	-	-
	-	*nipte*	*nipte*
WEDDING	*χolbohu:*	-	-
	-	*kurum*	*kurum*
NAPE OF NECK	*sürün:*	-	-
	-	*hergi*	*hergi*
CAMP	*tühülge*	-	-
	-	*hara:n*	*hara:n*
VALLEY	*χočo*	-	-
	-	*oŋχo*	*oŋχo*
PLAIN	*sïhï*	-	-
	-	*kitieme*	*kitieme*
IDOL	-	-	-
		hemeke:n	*hemeke:n*
SHEEP	-	-	-
		dʒollo	*dʒollo*

In these cases we can speak of full replacement, since the Sakha word is not remembered by current Dolgan speakers, if it were indeed inherited from Sakha, or does not exist. In the remaining cases the word from Evenki, which sometimes has an adjusted meaning, exists parallel to the word from Sakha, as is shown in the table below, which represents the clearest examples of this type.

Table 4.25: Copies from Evenki, Sakha word exists as well

Concept	Sakha	Dolgan	Evenki
MUSHROOM	*tellej*	*tellej*	-
	-	*dögömö:χtö*	*deginmekte*
BRAIN	*meji:*	*meji:*	-
	-	*irge*	*irge*
WHITE	*ürüŋ*	*ürüŋ*	-
	-	*če:lke:*	*čelke*
THE BOW	*ha:*	*ha:*	-
	-	*alaŋa*	*alaŋa*
BUTTOCKS	*emehe*	*emehe*	-
	-	*darama*	*darama*
NAKED	*hïgïnńaχ*	*hïgïnńaχ*	-
	-	*pelde:ki:n (?)*	*ńarbaki:n*

Tellej is used for mushroom in Sakha as well as in the variety of Dolgan spoken in the Anabar region, which is just across the border of the Taimyr and located in the Sakha Republic. The variety of Dolgan spoken in that area is more similar to Sakha than the Dolgan variety spoken in other areas. In the other Dolgan settlements *dögömö:χtö* is used, which is related to Evenki *deginmekte* (as a third alternative *kuna:χ* can be found, which also exists in Sakha). As far as I am aware, there is no semantic difference between those lexical items. As was mentioned in Section 4.4.2.1.1, *meji:* has undergone a semantic shift in Dolgan. For the concept BRAIN Sakha uses *meji:*, whereas Dolgan uses *irge* from Evenki. However, *meji:* still exists in Dolgan with the meaning of 'head'. The same holds for the concept WHITE, which is represented in Sakha by *ürüŋ* and in Dolgan by *če:lke:*, as in Evenki. However, *ürüŋ* is also still used in Dolgan, but with the more general meaning of 'light colour' or the 'light colour of reindeer fur'. 'Bow' is expressed in Sakha as *ha:*, and in Dolgan as *alaŋa*. *Ha:* exists in Dolgan with the meaning of '(unspecified) weapon'. *Emehe* exists in both languages with the meaning of 'buttocks' but Dolgan has an additional way of expressing this body part, which comes from Evenki. However, *darama* has undergone a semantic change from 'crotch' in Evenki to 'buttocks' in Dolgan. Finally, *hïgïnńaχ* exists in both languages with the meaning of 'naked', but Dolgan also has the word *pelde:ki:n* which could, according to Voronkin (1995), be related to Evenki *ńarbaki:n* 'naked'. The difference in meaning between

the two lexical items is that *hïgïnńaχ* in Dolgan is 'naked' but also 'too sparsely dressed', whereas *pelde:ki:n* has the meaning of being completely bare-skinned.

4.4.2.2.3 UNKNOWN ORIGIN

Besides the copies from Evenki or Russian discussed above, there are concepts that are expressed in Dolgan by lexical items that I was not able to trace back to Sakha, Evenki or any of the other neighbouring languages. This may be due to the fact that information in dictionaries is often incomplete, which would plausibly explain the absence of the words in Table 4.26 from the dictionaries at my disposal.

Table 4.26: Lexical items of unknown origin

Concept	Sakha	Dolgan
VAGINA	*abas*	*bökü*
PENIS	*übüs*	*öčö:*

The Sakha words *abas* and *übüs* (*öbüs* in Dolgan) also exist in Dolgan, but have a rude connotation according to my informants, whereas *bökü* and *öčö:* do not. Alternatively I may not have been able to find, or may not be sure of the related lexical item in other languages due to major changes in phonetic form or meaning, as I could imagine for *χapataj* in Dolgan, meaning 'bald', whereas Sakha uses *taraγaj*, or for *tömüje* for 'finger' where Sakha has *tarbaχ* and the neighbouring Samoyedic language Nganasan has *tɵrija*. Finally, some may just be language-internal innovations part of which could be motivated by onomatopoeic associations (e.g. *titire:, čapkahajdas*).

Table 4.27: Lexical items of unknown origin

Concept	Sakha	Dolgan
SHIVER	*ilibire:*	*titire:*
ENVY	*ïmsï:rï:*	*ordugurgo:*
SANDFLY	*oŋo:ju*	*kïrada:j*
SCREW	*bi:nte*	*hurbuk* (peg)

SPLASH	*ïhïaχtas*	*čapkahajdas*
STORM	*sillie*	*boloho*
PIECE	*to:roχoj*	*elteχ*

4.4.3 SUMMARY

The lexical comparison of Dolgan and Sakha has shown that these languages differ considerably in their lexicon (40.1% of the investigated meanings). However, a closer investigation has shown that a much smaller proportion of these differences can be attributed to contact with speakers of other languages. Since for many differences we cannot determine with certainty whether or not contact played a role in the change, the exact proportion of contact-induced change is hard to determine. To give an estimate, the number of copies from Russian (79) and Evenki (29) added to the probable cases of contact-induced semantic change discussed in this chapter (13) would add up to a proportion of approximately 15% (121:775). Regardless of this relatively low proportion, the character of certain differences is nevertheless indicative of a close relationship between Dolgans and Evenks in the past. This was exemplified by substance copies from Evenki in the domain of non-cultural vocabulary, and more significantly by the restructuring of semantic patterns in the Dolgan system of kinship terms. Other differences were explained through contact with Russians and for a minority of differences the motivations are unclear.

Section 4.4.1.1 showed that from a quantitative point of view the differences between Dolgan and Sakha are not restricted to particular semantic fields. This may indicate that changes are pervasive throughout the entire lexicon, but alternatively this could be due to the rather random allocation of certain concepts in the Loanword Typology list to a particular semantic field. Despite this pervasiveness, certain semantic fields show a higher percentage of differences than others, but the overall picture is a gradual cline rather than a striking pattern, in which only the semantic fields of 'the body' and 'kinship' stand out in their ranking from a cross-linguistic perspective.

Turning to the nature of these differences, the overwhelming majority was characterised as semantic changes (45.1%) followed by replacements (16.6%) and changes in form (15.6%). The category of semantic change is dominated by the type 'broader' (94.8%), where the Dolgan word covers a wider semantic area than

the same word does in Sakha. This type of difference pervades all semantic fields, except function words. Most of these semantic extensions seem to be language-internal developments, and only few can convincingly be argued to have been triggered by contact through comparison with the neighbouring languages. The emerging overall picture of vocabulary with less specific meanings, the relatively frequent use of descriptive phrases where Sakha has a single lexical item, loss of lexical items, and copies from Russian conspires towards the hypothesis that this generalising tendency in Dolgan is a recent development motivated by the ongoing shift to Russian.

However, other differences do point to contact. Within the semantic field 'the body' three semantic changes were argued to be motivated by contact with Evenki. These are the semantic extensions NOSE → BEAK, SOLE → FOOT, and the shift in HEAD → BRAIN. However, for these changes a language-internal motivation cannot be excluded, since the concepts in question are closely related and the direction of change follows cross-linguistic tendencies. The only exception here is the development from NOSE to BEAK, which would be expected to occur in the reverse direction from a cross-linguistic point of view. However, this would not demote contact as a potential explanation of this change. Language-internal and language-external factors are not mutually exclusive, they may have reinforced each other, and the cumulative effect may be reflected in these semantic changes.

Influence from Evenki is more difficult to deny in the explanation of the semantic restructuring of kinship terminology. First, the proportion of instances is higher (61.1%) and second, the restructuring of a kinship system reflects not only a linguistic change, but a more fundamental change in social organisation, which in turn can be explained through close contact and intermarriage between peoples with different social structure. The analysis has shown that Dolgan employs Sakha words for the concepts of BROTHER/SISTER, UNCLE/AUNT, and FATHER-IN-LAW/MOTHER-IN-LAW, but their semantic distribution matches the system of kinship terminology of Evenki.

With respect to foreign copies it appears that Russian copies are more common than copies from Evenki, and that are distributed over more semantic fields. The Russian copies were divided into four types, of which only non-cultural items such as 'mosquito' or 'swamp' could point to a closer relationship between Russians and Dolgans in the past, since they are not a corollary of the introduction of newly introduced concepts of modern society. In the responses to the concepts in the Loanword Typology list there are only six instances of this type, and

although they must certainly not be ignored, it would be overconfident to base bold statements on these few examples. Moreover, most of these Russian copies could be used in Sakha as well, albeit in some cases with a slightly different meaning. Hence, without evidence from historical texts, it is impossible to tell whether this is an instance of a single or two independent copying events. It could be the case that these terms were copied once into the ancestor language of Dolgan and Sakha, after which the meaning diverged in both languages, or it could reflect a situation where the variability in meaning is the result of the fact that the terms were copied into the languages independently. Russian copies of the other types could be due to more intense contact, but could equally well reflect the weaker position of the Dolgan language in the present, and thus be a more recent development.

Copies from Evenki are sparser than Russian copies both in their number and in their distribution. Compared to Sakha, Dolgan shows a higher number of Evenki copies for the investigated semantic fields, but a characterisation of the Dolgan language as Sakha with Evenki lexicon would by no means do justice to the actual facts.

4.5 INTERPRETATION OF RESULTS AND CONCLUSION

In order to formulate hypotheses about the sociolinguistic situation in which the current lexical differences between Dolgan and Sakha developed it is useful to present a schematic overview of the social settings and their expected linguistic outcomes for the different configurations of the Dolgan, Sakha, Evenk and Russian communities in different time frames. Since most contact that is relevant for the purpose of the reconstruction of the Dolgan history took place before Dolgan and Sakha were officially recognised as separate languages, in the schemes I will refer to the common ancestor language as Dolgan/Sakha, abbreviated as D/S.

As may be recalled from Chapter 3, the linguistic outcome of a contact situation is influenced by a complex interplay of linguistic and non-linguistic factors including 1) the structure of the languages in contact; 2) social dominance of the groups in contact; 3) linguistic dominance; 4) attitude and emblematicity. The relationships between the Dolgans and their neighbouring populations differ significantly with respect to these factors, depending on the neighbour in question (in particular the Sakha, Evenks and later the Russians), as well as the time period

during which the contact took place. Therefore it is important to take the various settings into account for the interpretation of the attested lexical differences. Each constellation is specified for the languages in contact, social dominance, linguistic dominance and the linguistic consequences this may have had for Dolgan/Sakha. These factors are considered from the perspective of both communities in contact.

Table 4.28 schematises the contact situation between Dolgans and Evenks until the 20th century, Table 4.29 between Dolgans and Russians during the pre-Soviet period, and Table 4.30 between the same groups after the institution of the Soviet regime. Table 4.31 represents the relation between Dolgans and Russians in most recent times. This division of the Russian contact into three time frames is necessary because the social and linguistic dominance relations were not comparable across these periods and had different linguistic effects on Dolgan.

Table 4.28: Contact situation between speakers of D/S and Evenki

	D/S perspective	Evenki perspective
Social dominance:	D/S	D/S
Linguistic dominance:	D/S	Evenki
Expected effect on Dolgan:	*Borrowing*: copies of Evenki cultural vocabulary into D/S	*Imposition*: imposition of semantic structure on D/S through L2 learners

While during the initial period of contact the Evenks may have been socially dominant because they occupied the area before the Turkic-speaking population arrived, these relations quickly changed judging by the establishment of the Khatanga Trading Way, which became associated mainly with the Dolgans, and where D/S became the lingua franca. It is this later configuration, which is represented in Table 4.28. According to this classification, the copies of cultural vocabulary would have entered D/S primarily through a process of borrowing, whereas semantic structures, including changes in kinship terminology, were imposed by Evenki speakers onto D/S. Within the growing community along the Khatanga Trading Way, where Dolgans, Russians and Evenks used to meet and to intermarry, Evenks who learned D/S as a second language may have introduced lexical items and other components from Evenki into their lect of D/S. The fact that such lexical items have become established in the Dolgan language of today could mean that the number of Evenks that learned D/S, and eventually shifted to it, was considerable.

Table 4.29: Pre-Soviet contact between speakers of D/S and Russian

	D/S perspective	Russian perspective
Social dominance:	D/S	D/S
Linguistic dominance:	D/S	Russian
Expected effect on Dolgan:	*Borrowing*: copies of Russian cultural vocabulary into D/S	

During the pre-Soviet period, the newly arrived speakers of D/S dominated the Taimyr socially as well as linguistically. This holds with respect to other native peoples as well as with regard to Russian settlers. We know that D/S was used as a lingua franca between indigenous people, and historical records report that Russians who came to live there during that time would also learn the language and after a while were 'indistinguishable' from the indigenous people. This suggests that the presence of these early Russian settlers did not change the relation of social or linguistic dominance, probably because they were not enough in number and because they were partially dependent on the native population for survival. Therefore Russian substrate effects as a result of imperfect learning are unlikely to have become established in the D/S language. The Russian lexical material that entered D/S during that time can be recognised as labels for unfamiliar cultural items (e.g. food, tools, etc.).

The situation for the Soviet period is different, as shown in Table 4.30.

Table 4.30: Soviet contact situation between speakers of Dolgan and Russian

	Dolgan perspective	Russian perspective
Social dominance:	Russian	Russian
Linguistic dominance:	Dolgan	Russian
Expected effect on Dolgan:	*Borrowing*: copies of Russian cultural and non-cultural vocabulary into Dolgan due to intense contact, cultural pressure and prestige of Russian	

During this time, Dolgan was already considered a separate language by some scholars, which is why in this table Dolgan is used instead of D/S. From the 1930's onwards, Russian influence, and social dominance, became more and more noticeable in Dolgan society, but until the 1970's most people remained linguistically dominant in Dolgan. However, due to intense contact, cultural pressure and increasing prestige of the Russian language, many cultural as well as non-cultural lexical items were borrowed into the Dolgan language.

The current situation is represented in Table 4.31.

Table 4.31: Current contact situation between speakers of Dolgan and Russian

	Dolgan perspective	Russian perspective
Social dominance:	Russian	Russian
Linguistic dominance:	Russian	Russian
Expected effect on Dolgan:	*Imposition*: introduction of lexical items from dominant Russian language onto non-dominant Dolgan	

The change in linguistic dominance from Dolgan to Russian took place gradually when people were forced to settle in villages and to go to boarding schools, where the use of Dolgan was forbidden. From then onwards, Russian became more and more socially as well as linguistically dominant, leading to more lexical changes as well as grammatical restructuring that is currently on-going as a result of imposition from their dominant Russian language onto their emblematic, but non-dominant Dolgan (see Chapters 7 and 8 for more changes due to contact with Russian).