

Chapter 8

The complexification of Tungusic interrogative systems

Andreas Hölzl

University of Zurich

The present study discusses the nature and development of a Tungusic phoneme **K-* that has proven difficult to reconstruct. It is only fully preserved in one sub-branch of the Tungusic language family and today is usually considered a velar fricative **x-* (Benzing 1956). However, there is evidence from a Tungusic language called Alchuka hitherto almost unknown outside China (Mu Yejun 1985, 1986, 1987, 1988). In this language the phoneme is, somewhat irregularly, preserved as unaspirated *k-*, which corroborates its reconstruction as a plosive (Rozycki 1993). The main focus of the paper is the role of the phoneme in the interrogative system of proto-Tungusic as well as the detrimental implications of its loss in most Tungusic languages. In proto-Tungusic the phoneme had the function of a *submorpheme* or *resonance* similar to English <wh> (Bickel & Nichols 2007, Mackenzie 2009). Its loss led to incoherent interrogative systems with a large number of individual forms that are synchronically opaque, i.e. to complexification. Finally, the question is addressed whether this can be considered a “change for the worse” as indicated by the title of this volume.

1 Introduction

Tungusic is a small and highly endangered language family scattered across Northeast Asia. According to Janhunen (2012), Tungusic can be divided into the Jurchenic (IV), Nanaic (III), Udegheic (II), and Ewenic (I) subbranches. The former two form the southern and the latter two the northern branch of Tungusic (see also Georg 2004). The enumeration of the individual subbranches with roman numerals follows Ikegami (1974). Altogether, Tungusic encompasses some twenty different languages.

This chapter investigates Tungusic interrogatives, also known as interrogative words, question words etc. Most Tungusic interrogatives started with an initial consonant for which several different reconstructions were proposed (see Rozycki 1993). For the time being, the uncertainty of this consonant will be indicated using the label *K-. The only subbranch that is known to preserve the feature in question is Nanaic, which consists of the languages Hezhen, Kilen, Kili, Nanai, Samar (Northern Nanai), Uilta, Ulcha, and Ussuri Nanai (e.g., Schmidt 1923, 1928, Alonso de la Fuente 2011, Janhunen 2012). For a first impression, consider Table 1 which lists four interrogatives from five different languages that represent all four Tungusic subbranches. Roman numerals refer to the subbranches of Tungusic. In Nanai, all four interrogatives start with an initial *x*-, while in Manchu, Udihe, or Oroqen no such phoneme is present. An underscore indicates the lack of the initial that was regularly lost in these languages. Unexpectedly, however, there is an initial *k*- in a language called Alchuka. This language is not very well-known, but offers some crucial information for our understanding of Jurchenic and Tungusic in general (Hölzl 2017).

Table 1: Some selected cognates of interrogatives in Manchu (Norman 2013), Alchuka (Mu Yejun 1986), Nanai (Ko & Yurn 2011), Udihe (Nikolaeva & Tolskaya 2001), and Nanmu Oroqen (Chaoke 2007). Not all variants are listed.

Meaning	Manchu IV	Alchuka IV	Nanai III	Udihe II	Oroqen I
how	-	-	xo:ni	_ono	_ooni
how many	_udu	kutu	xado	_adi	_adi
what	_ai	kai-	xai	_i:-	_i-
when	-	-	xa:li	_aali	_aala

Two of the interrogatives were lost in the highly innovative Jurchenic subbranch. Apart from regular phonological changes, some progressive (Udihe *ono*, Oroqen *aala*) and regressive vowel assimilations (Manchu *udu*, Alchuka *kutu*) have slightly distorted the picture. But there can be no doubt that these forms represent a valid set of cognates (cf. Benzing 1956: 114). The interrogatives form one coherent system with an initial *x*- in Nanai. This can be called a formal *resonance*, which will be indicated using a tilde, i.e. *x~* (Hölzl 2018b). This is meant to indicate its partial analyzability as opposed to clearly analyzable morphemes that are indicated with a hyphen, e.g. Nanai *xai-do* ‘what-DAT’. As Bickel & Nichols (2007: 209) put it: “parts of words resonate with each other and can therefore be extracted as meaningful formatives or morphemes.” They give the example of

English demonstratives that share an initial /ð/. In English a similar phenomenon is also known from certain lexical items such as *snore*, *sneeze*, *sniff*, *snuff* etc., all of which start with *sn~* and have a vague similarity in meaning. The concept of a resonance can also be usefully applied to interrogatives (Mackenzie 2009). The loss of the initial in Tungusic would be comparable to the loss of the <wh> in English, e.g. *(wh)at*, *(wh)en* etc. Usually, a resonance in an interrogative system is an indication of an old etymological connection. For instance, German *wo* ‘where’ and *warum* ‘why’ are synchronically unrelated, but share a resonance in *w~*. Historically, however, the first part of *war-um* (literally ‘where-around’) is a cognate of *wo*, which lost the final *-r*, unless it was followed by a vowel. In this case, *warum* represents a *fused* form that is no longer analyzable (Muysken & Smith 1990). In German, there are only two interrogatives without an initial *w~*, which is due to a prefix, i.e. *in-wie-fern*, *in-wie-weit* ‘how (far), to what extent’ (literally ‘in how far’). Tungusic languages only have suffixes and exhibit the word order Interrogative Noun (IntN). There is thus a natural tendency to build up resonances over the course of time. Rephrasing Givón (1971: 413), one might say that, in these languages, today’s resonance is yesterday’s morphosyntax.

The resonance in Tungusic is not as clear-cut as English initial /ð/ in demonstratives such as *this*, because unlike in English there are many other words with the same onset. However, the loss of the phoneme in lexical items such as ‘wind’ had no effect on their meaning, e.g. Manchu *_edun*, Nanai *xədun*, Udihe *_edi*, Nanmu Oroqen *_adin*. There simply is no set of weather-related words with an initial **K-*. On the contrary, the Tungusic phoneme had a certain functional load in the interrogative system (cf. Langacker 2001). This is why the loss of the phoneme led to a very incoherent interrogative system with a high number of individual forms that are synchronically unanalyzable. It changed the interrogatives in Table 1 from a so-called *fused* to an *opaque* system that no longer exhibits any signs of a former etymological connection (Muysken & Smith 1990). Of course, derivational or inflectional suffixes remained unaffected by this. For example, Manchu (IV) *ai-de*, Nanai (III) *xar-do*, Udihe (II) *i-du*, Chaoyang Oroqen (I) *i-du* ‘where’ (Hu Zengyi 2001: 261) are all still analyzable as ‘what-DAT’. Nevertheless, it will be argued in this chapter that the loss of the phoneme led to a *complexification* of the Tungusic interrogative system. In Manchu, for instance, there is no indication that *ai* ‘what’ and *udu* ‘how many’ could once have been related etymologically. §2 addresses the question how the complexity of interrogative systems can be described cross-linguistically.

One of the main problems for the comparison with Alchuka is the irregularity of the occurrence of the initial *k-*. For instance, it would be expected to be present in the word for ‘wind’ as well, but this word has been recorded as *_əduŋ*

(Mu Yejun 1985: 6). At least in some cases, the irregular correspondences can be explained as cases of borrowing from Manchu dialects that had a strong impact on Alchuka (cf. Aihui Manchu *_odon/ŋ*, Wang Qingfeng 2005: 162). For instance, the interrogative *kai-* has also been recorded as *_ei* by Mu Yejun (1986: 10). In this case, it is plausible to assume a certain amount of variation among what were probably the last speakers of the language when the data were collected during the 1960s (Mu Yejun 1985: 5). Again, the variant *_ei* could represent a loan from a Manchu dialect (cf. Yibuqi Manchu *_ei*, Zhao Jie 1989: 127). Furthermore, the recording of the Alchuka data by Mu Yejun is not always reliable. In specific cases, it is unclear whether the differences found in the recordings are true variation in the language itself or simple spelling mistakes. Nevertheless, independent evidence from other Jurchenic varieties confirms that the initial consonant as such is a real phenomenon and cannot be due to mere spelling mistakes (e.g., Kiyose 2000). §3 considers these problems relating to the Tungusic interrogative systems further.

The following research questions will be addressed in this study: (1) Is Alchuka *k-* a reflex of proto-Tungusic **K-*? (2) What, exactly, is the nature of the proto-Tungusic phoneme **K-*? (3) What were the consequences of its loss for the complexity of the Tungusic interrogative system? (4) How can the complexity of interrogative systems be defined cross-linguistically? (5) And finally, can this development be considered a “change for the worse” as indicated by the title of this volume?

The chapter has five subsections, including this introduction (§1). §2 briefly introduces seven different dimensions of complexity and defines the complexification of interrogative systems. Based on this general outline, §3 analyses the Tungusic interrogative system as well as its development through time. Given the rediscovered data from Alchuka, a new reconstruction for the proto-Tungusic phoneme **K-* is proposed. §4 evaluates the loss of the phoneme as well as its consequences and inquires whether it can be considered a change for the worse. §5 presents some conclusions.

2 The complexity of interrogative systems

Language is a very complex phenomenon. Every language requires a certain degree of complexity to be useful and engaging (cf. Norman 2010). In comparison, differences between languages are relatively minor, but nevertheless clearly observable. There are different approaches to complexity. For instance, Miestamo (2008) draws a distinction between absolute (i.e., objective) vs. relative (i.e., subjective) complexity on the one hand and between local vs. global complexity on

the other. Following Miestamo (2008), this study is only concerned with *absolute* and *local complexity*, i.e. it tries to objectively describe the complexity of one domain of the Tungusic languages. It will not evaluate the overall complexity of entire languages and will not refer to the relative difficulty in acquiring Tungusic interrogatives or the like.

Complexity can tentatively be defined “as the number of parts in a system or the length of its description” (Miestamo 2008: 27). However, complexity also depends on the quantity and quality of the interconnections of those parts (e.g., Karlsson et al. 2008: viii) and on the status of the system in its linguistic environment (e.g., Turvey 2009). The degree of complexity of a certain phenomenon can be described along different dimensions. For the purposes of this chapter, seven such dimensions will be differentiated (Table 2).

Table 2: Dimensions of complexity included into this study (loosely based on McWhorter 2007, Karlsson et al. 2008, Turvey 2009, Trudgill 2011).

Number	Dimension	Simple	Complex
1	regularity	regular	irregular
2	redundancy	underspecified	overspecified
3	analyzability	analyzable	unanalyzable
4	amount	few forms	many forms
5	organization	organized	unorganized
6	coherence	coherent	incoherent
7	delineation	bounded	unbounded

Depending on the domains of language under investigation, a different set of dimensions will be more adequate. The dimensions of organization, coherence, and delineation have been inspired by Turvey’s (2009) discussion of the notion of a nonsystem.

What can be meant by nonsystem? A set of isolated pieces that don’t interact, or interact so weakly that their influences upon each other are negligible, seems to fit the bill. Even better, perhaps, is the notion of a collection of related pieces where the relations have no implications for the properties or behaviors of the pieces. Certainly lacking in the image of a nonsystem is the sense of shared influences or mutual dependencies; intuitively, a nonsystem exhibits no coherence or functional unity. Also lacking is the sense of a boundary, a separation of the pieces into “ground” (pieces that surround) and “figure” (pieces that are surrounded). (Turvey 2009: 98f.)

Interrogatives cross-cut several different word classes (e.g., Dixon 2012: 409). For this reason, an interrogative system is usually difficult to present in one table. Depending on the language, some of the interrogatives can often be inflected according to their word class, while others cannot. In Manchu, for instance, *we* ‘who’ and *ai* ‘what’ can be inflected for case, while *udu* ‘how many’ can take morphology specific to numerals, e.g. *udu-ci* ‘how_many-ORD’, *udu-te* ‘how_many-DISTR’ etc. Nevertheless, despite their inherent heterogeneity, interrogatives can still be said to form one functional domain. If we take the concept of an interrogative system seriously (e.g., Muysken & Smith 1990), it should exhibit all the hallmarks of a system described by Turvey. The more it resembles a *nonsystem*, the more complex and unsystematic it is.

Diachronic changes in complexity can be called *simplification* and *complexification* (Trudgill 2011). This paper is a case study in complexification, i.e. the emergence of more complex structures out of simpler ones. In the following, the seven dimensions presented in Table 2 will be briefly defined and exemplified by Tungusic data. For a better understanding of the different dimensions, some cases of simplification will be mentioned as well.

2.1 Regularity

The first dimension of complexity is *regularity* (e.g., McWhorter 2007: 33–35; Trudgill 2011: 85ff.). Exceptions to rules or irregularities in language structure, such as suppletion, increase the number of elements the system has. An example of irregularity in the interrogative system can be found in several northern Tungusic languages. Most languages of this branch have a suffix that usually can only be encountered on one interrogative. In Udihe, for example, only the nominative or unmarked form of the interrogative *j’e-* exhibits the suffix *-u*. It is lacking on other interrogatives and is replaced if the interrogative is inflected for case, e.g. *j’e-u* ‘what’, *j’e-du* ‘what-DAT’ (Nikolaeva & Tolskaya 2001: 348). The suffix is also present in the closely related language Oroch. However, case markers regularly attach to the suffix, which in this case could be analyzed as an augmentation of the nominal stem instead, e.g. *jaa.v* ‘what’, *jaa.v-du* ‘what-DAT’ (Avrorin & Boldyrev 2001: 197). This is an example of a slight simplification of an interrogative system. In Manchu, there is no indication of the suffix left, e.g., *ya* ‘which’, *ya-de* ‘which-DAT’, i.e. the irregularity has been entirely lost.

Similar to cases of suppletion, exceptions from a resonance can also be considered an irregularity. The more exceptions from a resonance there are, the more irregular the interrogative system is. In Tungusic the two interrogatives **ŋüi* ‘who’ and **ja-* ‘(to do) what’ were the only exceptions from the resonance

in *K~. The interrogative *ja- has been entirely lost in Nanaic, which is also the only subbranch to have generally preserved the resonance. This can be considered a decrease in irregularity in Nanaic languages.

2.2 Redundancy

Redundancy has also been called overspecification (e.g., McWhorter 2007: 21–28). It is here understood as the number of different expressions for the same semantic category in a certain language has. Underspecification can lead to the creation of new forms and overspecification (or redundancy) can lead to competition and the loss of certain forms. A certain amount of redundancy must have already been present in the proto-Tungusic interrogative system. There seems to have been competition between two semantically very similar interrogatives that can roughly be reconstructed as *Kai- and *ja- ‘(to do) what, which’. These show an intriguing distribution among modern Tungusic languages (Hölzl 2018b: 315f.). The meaning ‘to do what’ is usually expressed by *Kai- in southern, but by *ja- in northern Tungusic. Ewenic and Jurchenic preserve both interrogatives, but Nanaic has completely lost *ja- and Udegheic has almost entirely lost *Kai-. It is only preserved in a few derived forms such as Udihe *i-du* ‘where’. In other words, both Nanaic and Udegheic have simplified their interrogative system by reducing the amount of redundancy. To this day, several languages preserve some of this redundancy in sometimes allowing both stems for the same derivations and inflections, e.g. Manchu (IV) *ai-de* vs. *ya-de*, Udihe (II) *i-du* vs. *j’e-du*, Even (I) *i-du* vs. *ja-du* ‘where’ etc. In some Ewenic languages the stems have partly merged phonologically, e.g. Khamnigan Evenki *i(i)-* vs. *i(e/i)-* (Janhunen 1991), which is another example for simplification.

The presence of several resonances, such as English /h/ (*who, how, how many/much*) and /w/ (in all remaining interrogatives), can also be considered a form of overspecification. If a resonance, such as English *w~* is in fact a submorpheme that carries a certain functional load (e.g., Langacker 2001, Mackenzie 2009), the existence of a second submorpheme *h~* with the same function must be considered redundant.¹

2.3 Analyzability

The dimension of *analyzability* is understood here in the following sense: The more analyzable a form is with the help of the other elements in a system, the less

¹This should not be confused with the first dimension that here includes exceptions from any resonance.

complex it is.² But analyzability is not always clear-cut. Instead, there is usually a scale of more and less analyzable forms. Partly analyzable forms, cranberry morphs, or resonances complicate matters considerably. Interrogatives in creole languages usually tend to be analyzable (Bickerton 2016 [1981]: 65; Muysken & Smith 1990: 884). However, an increase in analyzability is not necessarily restricted to creole languages. The Jurchenic branch of Tungusic, for example, is not a true creole but nevertheless exhibits some simplification due to non-native acquisition in its past (e.g., McWhorter 2007, Trudgill 2011, Hölzl 2018). Possibly, this can be observed in the interrogative system that exhibits a large amount of analyzable forms that are based on the interrogatives *ai* and *ya*, e.g. Manchu *ai-ba-*, *ya-ba-* ‘where’ (*ba* ‘place’) etc. If no new forms are created, analyzability tends to decrease over the course of time. For instance, Manchu *ai-ba-de* ‘what-place-DAT’ has a variant *ai-bi-de* ‘what-?place-DAT’. The second element is a cranberry morph that most likely derives from *ba* ‘place’. There is even a less analyzable variant *abi-de* that must be the result of an additional contraction (e.g., Norman 2013). Both *-bi-* and *a-* are no longer clearly analyzable within the system and therefore increase the number of elements. The suffix *-u* in Udihe only occurs on one interrogative and therefore is not analyzable with the help of other elements within the system.

From the perspective of analyzability, a resonance can be viewed as making a system more and less complex at the same time. On the one hand, a resonance by definition is only partly analyzable and therefore, makes individual forms with the resonance more complex. On the other hand, it allows at least a partial analysis of all forms that exhibit the similarity. The loss of a resonance leads to the loss of an etymological connection between the individual interrogatives and thus to a decrease in analyzability. Due to this loss of analyzability, the number of individual interrogatives rises considerably.

2.4 Amount

The number or *amount* of interrogatives appears relatively straightforward: the more interrogatives a language has, the more complex it is. However, the number depends on the analysis. One possibility would be to include “basic” interrogatives, exclusively (Hengeveld et al. 2012). However, this approach suffers from the problems of analyzability mentioned above. Among Tungusic languages, the lowest number of basic interrogatives seems to be present in Udegheic. In Udihe, for instance, there are only four unanalyzable forms: *ni(:)* ‘who’, *ali* ‘when’, *adi*

²Notice that this dimension does not take into account the number of elements an analyzable form exhibits.

‘how many’, and *ono* ‘how’ (Nikolaeva & Tolskaya 2001). As we have seen above, even *j’e-u* ‘what’ contains a suffix. Of course, Udihe has a wealth of additional interrogatives, but all of them are analyzable to different degrees. Most of them are derivations and inflected forms of *j’e-* ‘(to do) what’. Some of those have a parallel based on the stem *i:-* that is only preserved in derivations (1).

- (1) Some variants of Udihe interrogatives (Nikolaeva & Tolskaya 2001)
j’e-du, i:-du ‘what/which-DAT > why, where’
j’e-le, i:-le ‘what/which-LOC > where’
j’e-mi, i:-mi ‘what/which-CVB > why’

The forms make the impression of mere variants and are perhaps perceived as such by the speakers, but they are etymologically distinct (**ja-* vs. **Kai-*). In other words, the forms based on *i:-* are synchronically only partly analyzable, because they contain some form of cranberry morph. It is an open question whether such partly analyzable forms should be treated as “basic question words” or not.

The same problem applies to interrogative systems with a resonance. If, for example, a resonance is counted as one interrogative stem, Tungusic most likely had only three interrogatives, i.e. **ñüi* ‘who’, **ja-* ‘(to do) what, which’, and **K~* (e.g., Benzing 1956; Hölzl 2018b: 312–330). If, on the other hand, forms with a resonance that are otherwise unanalyzable are counted as well, their number increases substantially. If the loss of a resonance is a decrease in analyzability, theoretically the number of “basic” interrogatives should rise, too.

2.5 Organization

Languages differ in their overall *organization* of the interrogative system. In Tungusic, for example, there is no special interrogative meaning ‘where’. Instead, all languages employ case-marked interrogatives meaning ‘what’ or ‘which’ to express that notion. In these languages the interrogative meaning ‘where’ is simply part of a paradigm. The inclusion into a paradigm could be interpreted as a form of regularity. However, this kind of organization leads to anomalous case forms that differ semantically from the rest of the paradigm, e.g. Udihe *j’e-we* ‘what (ACC)’, but *j’e-du* ‘where’. From this perspective, an interrogative system with a special locative interrogative, such as English *where*, could be considered more organized.

The dimension of organization can perhaps be applied to the semantic scope of a resonance. Ideally, the semantic scope of an interrogative covers a coherent region in semantic space (on which see Cysouw 2005, 2007; Hölzl 2018b: 82f.

and references therein). More research is necessary on whether exactly the same principles can be applied to resonances. But because resonances usually emerge through the spread of one interrogative over several semantic categories and its subsequent decrease in analyzability, this is a plausible scenario. For instance, cross-linguistic research seems to indicate that the two categories *THING* ('what') and *QUANTITY* ('how many/much') can only be expressed by the same form if *MANNER* ('how') is also expressed in the same way (e.g., Cysouw 2005). However, in Kilen the resonance $\chi\sim$ can only be found on χai 'what' and χadu 'how many', but not on $_oni$ 'how' (An Jun 1986). This unorganized system is the result of language contact in which the form $_oni$ was borrowed from a northern Tungusic language. Nanai still has a more organized system with the form $xo:ni$ instead (§3).

2.6 Coherence

Without doubt, *coherence* (what holds the system together) is the most important dimension for this paper. There are several possibilities, but this study describes the coherence of an interrogative system in terms of resonances (i.e., formal coherence). One of the most striking examples of complexification in Tungusic interrogatives was triggered by the phonological change pointed out in §1. Consider Tables 3 and 4, which list Uilta and Evenki interrogatives as examples of coherent and incoherent systems, respectively. Uilta ηui and Evenki $ni:/nɪ$: both go back to $*\eta\ddot{u}i$. Tungusic $*ja-$ has disappeared without a trace in Uilta, but is still present in Evenki as $\text{æ}:-$. Finally, the resonance $*K\sim$ is preserved in Uilta as $x\sim$ but has been lost in Evenki. Interestingly, Evenki has many forms starting with $i(:)/ɪ(:)\sim$, while no such vowel follows the initial in Uilta (see §3 for implications). This could be considered a secondary resonance in $I\sim$ that has been built up following the loss of the original resonance. However, there are several other forms such as $\text{ɔ}:\text{qin}/\text{ɔ}:\text{kin}$ 'when' that do not conform to this pattern.

An important problem for this study is the question whether the presence of a *resonance* is a complicating or simplifying factor. Perhaps, from the point of view of analyzability alone, a resonance makes things more complex by being only partly analyzable. From the point of view of coherence, however, one homogenous resonance as in Uilta (Table 3) could be said to be a simplifying factor instead because it holds the system together.

2.7 Delineation

The dimension of *delineation* (what differentiates the system from other elements) refers to the status of the interrogative system in a given language. A

Table 3: The formally coherent interrogative system of Uilta (III) (Ikegami 1997).

Form	Meaning
ɲui	who
xaali	when
xaawu, xauwu	which one
xai(-)	(to do) what
xaidu	where
xaimi	why
xamačču	whence
xamačiga	what kind of
xasu	how many/much
xawasai	whither
xawwee	where, what place
xooni	how

Table 4: The formally incoherent interrogative system of Aoluguya Evenki (I) (Hasibate'er 2016: 171, 238).

Form	Meaning
ni:, nɪ:	who
æ:qɔn, æ:ɣɔn	what
æ:χa	how
adɪ, addi	how many
ɔ:qm, ɔɣ:m	when
irəgeçin, irgətʃin	what kind
i:du	where
i:li	where
i:rba	how much
i:ʃ	which
ira:	which one
irga:	how much
i:la, i:ra	where

bounded system might exhibit certain phonological or morphological properties that are not found outside of the system. In English, for instance, an initial /ð/ is almost exclusively encountered in the demonstrative system (Bickel & Nichols 2007: 209). Perhaps an analogy from visual perception can help make this point even clearer (cf. Turvey 2009). A monochromatic piece of paper (a coherent interrogative system with one resonance) can be perceived much better than a multi-coloured one if they are held up before of a heterogenous background (the linguistic system). Of course, it can be perceived even more clearly if the colour is not found in the background at all (if the resonance is restricted to the interrogative system). An example can be found in the nearby Turkic language family. It has long been noted that an initial *n*- in proto-Turkic was restricted to the interrogatives (see Hölzl 2018b: 354 and references therein). To my knowledge, no comparable phenomenon is known from Tungusic interrogatives. Similar to English /w/ or /h/ in the interrogatives, the Tungusic initial **K*- also occurred on several other words.

However, there are some examples of the inflectional delineation of a interrogative system. In the Ewenic language Even, for instance, there is a suffix *-k*, which is cognate with Udihe *-u*, Oroch *-v*, and Evenki *-qɔn/-ɣɔn* encountered be-

fore. Unlike these languages, Even *-k* can also be found on another interrogative and two demonstratives (Benzing 1955: 77, 79). In this case, there is coherence in a subset of the interrogatives. Often, the interrogative system is only weakly delineated from the demonstrative system (Diessel 2003). The two systems tend to have a certain amount of parallels and overlap in inflection or derivation. There can also be a formal resonance between interrogatives and demonstratives as a result of this, e.g. English *whither*, *hither*, and *thither* etc.

The different dimensions of complexity tentatively proposed in this section show complex patterns of interaction. For instance, the lack of analyzability is not only correlated with a higher number of individual forms, but also with incoherence. In Turvey's (2009: 99) terms, the lack of analyzability with the help of elements within the system leads to a lack of "mutual dependencies" and therefore, to less coherence. Another example of such an interaction exists between coherence and delineation. A formally coherent interrogative system with one resonance in all interrogatives is more easily delineated from the rest of the language than one without any coherence.

3 Loss of the resonance in Tungusic interrogatives

The phenomenon investigated in this section is the loss of the proto-Tungusic phoneme **K-* in word initial position.³ As pointed out in §1, this regular phonological process was extremely detrimental to the interrogative system in most languages, where it fulfilled the role of a *submorpheme* similar to English <wh> (Bickel & Nichols 2007, Mackenzie 2009). The implications of the loss will be pointed out in §4. Tungusic is one of several language families in Northeast Asia (NEA) and surrounding regions to exhibit what has been called *K-interrogatives*: more than two interrogatives in a given language start with the same velar or uvular plosive or fricative (Hölzl 2018b: 6, 405f., 432). Other language families with this feature include, for example, Mongolic, and Turkic. As pointed out in §1, a resonance usually indicates an etymological connection. A similar resonance across different language families is first and foremost a typological similarity, but could also indicate a certain connection in terms of language contact and/or a genetic relatedness. More research is necessary on their global distribution and origin, but *K-interrogatives* appear to be a relatively stable phenomenon in NEA and their loss in Tungusic is rare, if not unique.

³It may also have existed in word-internal position, see Janhunen (2017) and references therein for some discussion. For reasons of space, this question cannot be addressed here.

Traditionally, it was believed that the phoneme **K-* was lost everywhere but in Nanaic (e.g., Benzing 1956: 41f.). There is, however, also some evidence that the phoneme may have been present in Ewenic at some point in time, where very few isolated relics with an initial *h-* were preserved in peripheral varieties (e.g., Vasilevich 1958; Doerfer 1973: 581), e.g. Sakhalin Evenki *_ure* ‘mountain’ (Nanai *xurān*), but *herekī* ‘frog’ (Nanai *xəra*) (Bulatova & Cotrozzi 2004: 106f.). It has also been speculated that a form of Jurchen that can be called Jurchen A (Grube 1896, Kiyose 1977) may have a few forms with an initial **h-* as well (Kiyose 1996, 2000, Hölzl 2017). Certain modern Jurchenic varieties potentially also preserve an initial *h-* in some relics, e.g. Written Manchu *_amaha*, but *hamuha* ‘afterwards, later, future’ as recorded in Qitamuzhen (Nanai *xama-*, dial. Evenki *hama-*) (Hölzl & Hölzl 2019). However, it was previously not widely known that the initial may also have been preserved as *k-* in yet another language from the Jurchenic branch called Alchuka (Hölzl 2017, 2018b). If correct, this clearly demonstrates that the phoneme was present in proto-Tungusic and must have been lost at a later stage. Additionally, this could give additional evidence for the primary split of Tungusic into northern and southern Tungusic as proposed by Georg (2004) or Janhunen (2012). Most likely, **K-* was generally lost in northern Tungusic – there are only a few relics in Ewenic and none in Udegheic –, but was preserved in southern Tungusic. It could have been lost at a relatively late stage in the majority of Jurchenic.⁴

There have been several different more specific reconstructions of the phoneme **K-*, the most important of which are collected in Table 5. The difficulty of the reconstruction is due to the fact that the phoneme is only fully preserved in Nanaic. The newly found data from a Jurchenic language can potentially contribute much needed information for its reconstruction.

Most of the reconstructions are rather problematic and contradict what is known from cross-linguistic research on language change such as the cline in (2).

$$(2) \quad k > (kx >) x > h > \emptyset$$

As Bybee (2015: 29) points out: “These paths are unidirectional; that is, the changes always proceed from stop to affricate to fricative to /h/ to zero, and not in the other direction.” There is no evidence for an affricate in any Tungusic language (cf. Cincius 1949: 250), which has been put into parentheses. This general tendency also contradicts Doerfer’s (1973) assumption of a change of Tungusic **h-*

⁴Udegheic seems to have a few cases that were borrowed from Nanaic, e.g. Oroch *xuju(n)* ‘nine’ (Ulcha *xuju(n)*, cf. Udihe *je(j)i*).

Table 5: A summary of the most important previous reconstructions of Tungusic *K-.

Source	Reconstr.	Description
Schmidt 1923: 232	*x-	voiceless velar fricative
Shirokogoroff 1931: 244f.	*Ø	later prothetic development
Cincius 1949: 250	*kxh-	aspirated voiceless velar affricate
Benzing 1956: 41ff.	*x-	voiceless velar fricative
Doerfer 1973: 579–591	*h-	voiceless glottal fricative
Cincius 1975: 300	*k’-	voiceless palatal plosive
Rozycki 1993: 211	*k’-	(un)aspirated voiceless (velar) plosive

to Nanaic *x-*. If Alchuka *k-* can be shown to be an actual reflex of proto-Tungusic *K-, the reconstruction would have to be changed to a plosive as well. Rozycki (1993), based on external comparisons, has also quite convincingly argued for the reconstruction as a plosive.⁵ In addition, there are several areal parallels for a change from a velar plosive to a fricative in the interrogative system of, for example, Turkic and Mongolic languages (Table 6).

In order to better decide which, if any, of the reconstructions is the most adequate, the actual reflexes observed among Tungusic languages have to be consulted. Table 7 represents a part of the consonant inventory of the *International Phonetic Alphabet* (IPA). All attested reflexes in modern Tungusic languages are printed in boldface. As can be seen, there is a wide variety of different reflexes that include both one plosive (i.e., [k]) and six different fricatives differentiated by their place of articulation (i.e., [s], [ʃ], [ç], [x], [χ], and [h]). All fricatives and perhaps also the plosive are voiceless, which must be a feature inherited from the proto-Tungusic phoneme.

The sounds mentioned in Table 7 are a summary of the entire language family and cannot all be found in one single language. The plosive is only sufficiently attested in Alchuka. However, a Jurchenic variety that I call Chinese Kyakala potentially also has one example of an initial *k-*, i.e. Manchu *urun*, Chinese Kyakala *kulun* (or perhaps *kurun*) ‘wife, bride’ (see Hölzl 2018a, Hölzl & Hölzl 2019). Some peripheral Ewenic and perhaps Jurchenic languages exhibit an *h-* (Doerfer 1973, Kiyose 1996, 2000, Hölzl 2017). All fricatives are otherwise only attested in Nanaic.

⁵Please note that this study is mostly based on data from Tungusic languages. For reasons of space, external comparisons, such as with Mongolic languages, will be mentioned only briefly (e.g., Doerfer 1985, Rozycki 1993, Janhunen 2017 and references therein).

Table 6: Areal parallels for the lenition from plosive to fricative in the interrogative system (Hu Zhenhua & Imart 1987, Anderson 1998, Yamakoshi 2007, 2011).

Turkic	Fuyu Kirghiz	Khakas
what kind of	ɢadah, ɢadɪh	xayday
when	ɢajan	xaʒan
where	ɢayda	xayda
which	ɢayzɪ	xayzɪ
Mongolic	Khamnigan Mongol	Shineken Buryat
how many	kədui	xedii
when	kəzie	xezee
where	kaa-	xaa-
who	kən	xen

Table 7: IPA symbols for the phonetic space in question (voiceless / voiced). Attested reflexes of **K*, including allophones but not \emptyset , are in boldface.

	Alv.	Postalv.	Alv.-pal.	Retr.	Pal.	Vel.	Uvul.	Phar.	Glott.
Plosive	t / d	–	–	ʈ / ɖ	c / ɟ	k / g	q / ɢ	–	ʔ / –
Fricative	s / z	ʃ / ʒ	ɕ / ʑ	ʂ / ʐ	ç / ʝ	x / ɣ	χ / ʁ	ħ / ʕ	h / ɦ

For example, there are three different reflexes in Hezhen, which seems to be the maximum among Tungusic languages. Some exceptions apart, the nature of the sound can be predicted by the following vowel.

(3) Reflexes of **K*- in Hezhen (An Jun 1986: 79f.)

**K* > ɕ- | _i

**K* > x- | _ə, _u, (_i)

**K* > χ- | _a, _o

The set of reflexes in Table 7 differs significantly from that proposed in Benzing (1956: 41), who, apart from \emptyset , only mentioned *s*-, *x*-, *h*-, and, problematically, *n*-. The nasal appears to be a mistake that resulted from a misunderstanding of a secondary innovation in Manchu. Benzing (1956: 43) mentions the two Manchu examples (*n*)*imenggi* ‘oil’ and *nimanggi* ‘snow’ that, the differences in derivational

suffixes apart, correspond to, for example, Uilta *simuksə* and *simana*, respectively (Ikegami 1997). However, the correspondence of *n-* and *s-* is only valid at a first glance. Consider the comparison in Table 8. There is a relatively clear correspondence between *n-* in Manchu and \emptyset in Alchuka, especially with a following *m*.⁶ Crucially, this is a later phenomenon that can also be found in loanwords such as *niman* ‘goat’ that do not have a Tungusic background, but derive from surrounding languages such as Khitan (see Tang 2011). In other words, the initial *n-* in Manchu cannot be a reflex of Tungusic **K-*. The presence of the initial *n-* in *(n)imenggi* ‘oil’ and *nimanggi* ‘snow’ must be considered a coincidence (cf. Hezhen *_imaɣa* ‘fish’ etc.).

Table 8: A comparison of Manchu and Alchuka (Mu Yejun 1985, 1987, Norman 2013). JA = Jurchen A (Kiyose 1977).

	Manchu	Alchuka
fish	nimaha	_imaha
goat	niman	_iman
mulberry tree	(n)imala(n)	_imala
oil	(n)imenggi	JA *_imengi
snow	nimanggi	_imaŋi

Concerning the reflexes of **K-* in Nanaic, consider Table 9. Only a selection of examples and sources available for Nanaic languages was chosen. The primary split of the phoneme in Nanaic appears to have been based on the following vowel. As seen for Hezhen above (3), the reflex usually is an *s*-like sound in front of *i* (or *ɪ*) and an *x*-like sound otherwise (e.g., Benzing 1956: 41f.). There are some language-specific problems that cannot all be addressed here. For instance, Tsumagari (2009: 2) notes that an */s/* in Uilta is only realized as [s] before the vowels *a* and *o* [ɔ]. Before all other vowels, including *i*, it is pronounced as a [ʃ] or [sʃ].

Especially older descriptions suffer from an unclear and inexact notation of phonemes. It is not entirely clear, for instance, what sound the initial <ch> in Kilen mentioned by Jettmar (1937) in his German description represents. In German, a <ch> would normally be pronounced as [ç] before an *e*, but it is doubtful that this rule should apply here. Most likely, it represents a [x] instead, which is another allophone of <ch> in German. This is one of several examples where a

⁶There are, however, several irregularities regarding the initial (palatal) nasal in Jurchenic and Kilen that deserve a treatment of their own.

Table 9: Reflexes of **K*- in Nanaic according to several different authors in alphabetical and chronological order. Inner-Tungusic loanwords are given in parentheses. Not all variants are mentioned. Accents removed.

Language	back(wards)	wind	snow	Source
Hezhen	hami(kə)	hət̚ʃ	himana, ʃimana	Ling 1934
	χamilə	xədun	ximanə	An Jun 1986
Kilen	(_amidzikə)	hət̚ʃ	(_imana)	Ling 1934
	?	?	chemana	Jettmar 1937
	(_amidzgə)	(_ədin)	(_imanə)	An Jun 1986
Kili	(_amaski)	(_ədi ⁿ)	(_emana)	Sunik 1958
Nanai	xamasi	ʃədun	ximana, simota	Grube 1900
	xamasi	xödun [-ə-]	ximana	Schmidt 1923
	xamasi	xədun	simana, simata	Ko & Yurn 2011
Samar	?	xödu(n) [-ə-]	simana	Schmidt 1928
Ulcha	xamasi	xödu [-ə-]	simata	Schmidt 1923
	xamasi	xydu	xemana, simata	Majewicz 2011
Uilta	hamasai	huidö [-ə]	simana, simatta	Nakanome 1928
	xamaša	xydu	simani, simat(t)a	Majewicz 2011
	xamasai	xədu	simana, simatta	Ikegami 1997
U. Nanai	hamela	hedou [-u]	?	Venjukov 1862
	χamas'ɪ	xədu(n-)	s'im(a)na, s'im(a)ta	Sem 1976

velar-like fricative is preserved in the word for snow. This, as well as the complementary distribution of the s-like and x-like phonemes, are the main arguments for the assumption that the same phoneme **K*- was present in this word and in similar cases.

Table 10 lists the Nanaic cognates of three interrogatives according to the same sources as in Table 9. Given that in no Nanaic interrogative the resonance was followed by an *i* or *ɪ*, the velar-like phoneme is preserved everywhere (cf. §2).

The question whether Alchuka *k*- is a reflex of Tungusic **K*- is extremely complex and difficult to answer. Not all problems can be solved or even addressed in this chapter. As mentioned in §1, there are certain irregularities. Table 11 lists all attested interrogatives in Alchuka. Apparently, **ja*- has not been recorded. Most likely, *p'ə* derives from **ŋüi*, but this cannot be a regular continuation (Hölzl 2018b: 314). The resonance is only present in five out of the ten remaining recorded interrogatives. Those without the initial might represent borrowings from Manchu dialects. However, only Bala has an *n* in the word for 'when' (Hölzl 2018b: 330).

Table 10: Reflexes of **K*- in Nanaic interrogatives according to several different authors in alphabetical and chronological order. Likely inner-Tungusic loanwords are given in parentheses. Accents removed.

Language	what	how many	how	Source
Hezhen	hai	?hadu	?	Ling 1934
	?	?	?	An Jun 1986
Kilen	hai	hadu	hōni-	Ling 1934
	?	(_adi)	?	Jettmar 1937
	χai	χadu	(_oni)	An Jun 1986
Kili	(_ii-)	(_adi)	(_ōni)	Sunik 1958
Nanai	xai, hai-	xadu, hadu	xoń(e)	Grube 1900
	xai	xadu	xoŋe	Schmidt 1923
	xar	xado	xo:ni	Ko & Yurn 2011
Samar	xai	?	?	Schmidt 1928
Ulcha	xai	xadu	xōni	Schmidt 1923
	xaj	?xadum	xon(i)	Majewicz 2011
Uilta	hai	-	hōni	Nakanome 1928
	xaj	-	xōni	Majewicz 2011
	xai	-	xooni	Ikegami 1997
U. Nanai	hai	?	honi	Venjukov 1862
	χar	χado, χadʊ	χon'(i)	Sem 1976

In general, it is possible to identify several different categories. First, there are words with an initial *k*- that have a clear correspondence in Nanaic. Second, there are words with an initial *k*- that do not have a correspondence in Nanaic. Third, there are many words that would be expected to exhibit the initial *k*- based on Nanaic data, but do not. Fourth, in a few cases there is a potential external comparison outside of Tungusic (on which see also Rozycki 1993).⁷ Table 12 mentions three examples of each category. Finally, there are at least two cases in which the initial *k*- has comparisons in Jurchenic (Hölzl 2017). The list is not exhaustive, but sufficient for the purposes of this paper.

Theoretically, the initial *k*- in Alchuka could be a later prothetic development that is specific to this language (cf. Shirokogoroff 1931). Given the strongly suffixing character of all of Tungusic, it is implausible to assume an otherwise un-

⁷Potentially, some of the interrogatives in Tungusic could have a Mongolic origin, too, but this requires further research.

Table 11: Interrogatives in Alchuka (Mu Yejun 1986, 1987, 1988) in comparison with Manchu (Norman 2013). Likely inner-Tungusic loanwords are given in parentheses (Hölzl 2018b: 317).

	Alchuka	Manchu
who	ʔp'ə	we
for what reason	(_ei) t'uku	_ai turgun
how	katiram	_adaramē
how many	kutu	_udu
to do what	kai-na-mei	_ai-na-mbi
what	(_ei)	_ai
what has happened	gai-na-hanbie	_ai-na-habi
what (is it)	kent'aka	_antaka
when	(_ant'aŋgi)	_atanggi
where	(_ai-və-t)	_ai-ba-de
why	(_einu)	_ainu

known prefix *k-*. One should not exclude the possibility of a prothetic development for some cases, especially those of category two that have no correspondence in Nanaic. However, there is evidence that at least in some cases the *k-* cannot be a secondary innovation. Given the fact that there are Nanaic correspondences in category one, the problem cannot be due to chance. These examples cannot be explained by borrowing from Nanaic either. For example, Alchuka *kutu* contains a vowel assimilation specific to Jurchenic and *kai-na-* has a verbalizer that does not occur in this form in Nanaic. As indicated in §1, at least some examples of the third category can be readily explained with borrowing from Manchu dialects, which appear to have had a strong influence on Alchuka. In many cases, this might explain the absence of the initial *k-* that would otherwise be expected on the basis of a comparison with Nanaic. This is especially plausible if there are doublets such as *kai-* vs. *_ei* ‘what’. These must reflect an autochthonous and a borrowed form, respectively. It should be noted that the same problem exists for the three Nanaic languages Kili, Kilen, and Ussuri Nanai, which have many loanwords from Ewenic, Udegheic, and Jurchenic without the initial. This explanation is especially convincing if a given loanword exhibits additional features that are only attested in another language. For instance, Kili *_ədiⁿ* ‘wind’ not only lacks the initial consonant that is present, for example, in Uilta *xədu*, but the vowel *i* in the second syllable is a feature specific to northern Tungusic (Benzing 1956). In the case of Alchuka, such identifying features

Table 12: A comparison of Manchu (Norman 2013), Alchuka (Mu Yejun 1985, 1986), and Uilta (Ikegami 1997). JA = Jurchen A, JM = Jing Manchu, Kh. = Khalaj, MK = Middle Korean, PM = proto-Mongolic. Not all variants attested for Alchuka are shown.

Category	Meaning	Manchu IV	Alchuka IV	Uilta III
1	how many	_udu	kutu	xadu
	twenty	_orin	(k)ɔrin	xori
	what	_ai-	(k)ai-	xai-
2	this	_e-re	kə-r(ə)	_ə-ri
	this way, here	_ebsi	ke'uzi	_əwəsəi
	to become	_o-	(k)ɔ-	_o-
3	nine	_uyun	_ujen	xuju
	what	_ai	_ei	xai
	wind	_edun	_ɔduŋ	xədu
4	twenty	_orin	(k)ɔrin	PM *kori/n
	virtue	_erdemu	kərdem	Kh. här 'man'
	-Q	_o	(k)ɔ	MK (k)o
5	nineteen	JA *_onioxon	(k)uniku	JM kuniu
	to meet	_aca-	katʃ'a-	Bala hatʃ'a-

are often difficult to find because all languages involved are relatively closely related. More research on Manchu dialects is necessary in order to identify the exact source of the borrowings.

Potentially, some of the words with an initial *k* could also represent spelling mistakes that are not uncommon in Mu Yejun's data. However, one should not jump to the conclusion that all of the examples can be explained in this way. For example, the initial *k*- in numeral nineteen has been independently confirmed by Aixinjueluo Yingsheng, who remembered to have heard the form *kuniu* in his youth (see Aixinjueluo Yingsheng 2014, Hölzl 2017).

Apart from the comparison with Nanaic, there are additional indications that the initial *k*- in Alchuka is neither due to chance, nor a spelling mistake. Crucially, there are a few potential comparisons outside of Tungusic that deserve further discussion. For instance, the question marker =o in Manchu that lacks a Tungusic background most likely is a loan from Middle Korean -(k)o (Hölzl 2018b: 213). Furthermore, there is one example (*katʃ'a*- 'to meet') that allows a comparison

with an initial *h*- in Bala *hatʃa*- and Jurchen A **hača*-. Problematically, this initial *h*- is similarly irregular (e.g., Jurchen A **_onioxon* ‘19’) and is only attested in a few words (Kiyose 1996, 2000, Hölzl 2017, Hölzl & Hölzl 2019). But it represents additional evidence that the *k*- in Alchuka is neither an isolated phenomenon, nor a spelling mistake. The initial *h*- in Evenki dialects is similarly problematic but is still accepted as a valid correspondence by Doerfer (1973).

Yet another problem concerns the nature of the phoneme in Alchuka. It is usually written as <*k*> in Mu Yejun (1986), but as <*g*> in Mu Yejun (1987, 1988) (see Table 11). Descriptions of Jurchenic varieties disagree on the nature of the plosives. More research on the phonology of Jurchenic is necessary to determine the exact phonetic value of the plosives. It is possible that, at least in some varieties and similar to Mandarin (e.g., Zhao Jie 1989), the distinction between <*g*> and <*k*> is only one of aspiration ([*k*], [*kʰ*]) and not of voice as well ([*g*], [*kʰ*]). But Norman (2004: 27) argues that, in Manchu, a <*g*> is only pronounced as a voiceless unaspirated [*k*] in initial position. In any case, Alchuka <*k*> (Mu Yejun 1986) also corresponds to what is usually considered a voiced velar plosive <*g*> in Manchu. If Alchuka *k*- is indeed a reflex of Tungusic **K*-, it must have historically merged with the reflex of the original **g*-. Interestingly, the irregularity in Alchuka seems to include both the reflexes of **K*- and **g*-. For instance, the interrogative *_ei* is attested in the complex expression *ei əl'un ə'ɔ* (Mu Yejun 1986: 10), a cognate of Manchu *ai gelhun akū* ‘how dare ...’ (Norman 2013). The lack of several word-internal consonants is a different problem. But the cognate in Alchuka also lacks an initial *k*- that would be expected in *_əl'un* ‘timid’. Thus, it seems that the question of the initial *k*- in Alchuka is a more general problem. Future research will have to explain the sporadic loss of the initial **g*- and some other consonants, which goes beyond the possibilities of this study.

Based on the evidence in this section, a more detailed reconstruction of **K*- might be possible. The Alchuka data apart, the reconstruction as **x*- is, of course, very convincing, because it fits very well into the proto-Tungusic consonant system and also has a potential areal parallel in Mongolic (Janhunen 2017). However, as seen above, the phoneme could well have been a plosive rather than a fricative. A crucial question is the general structure of the proto-Tungusic obstruent system (Rozycki 1993, Janhunen 2017). According to the traditional reconstruction (Benzing 1956: 27), Tungusic had the velar consonants **g* and **k* (i.e., [*g*], [*kʰ*]). In most languages, **K* shows a different set of reflexes than **g* and **k*. Consequently, it must have differed in some respect from the other velar plosives. Rozycki (1993: 211) assumes that there might have been a distinction in aspiration (i.e., [*k*], [*g*], [*kʰ*]), and indeed the Alchuka data potentially give additional

evidence for this point of view. However, there are several additional possibilities such as a difference in the place of articulation instead of the manner of articulation. For example, several languages have an alveolar-palatal or uvular reflex of the phoneme *K-, which suggests that it could theoretically also have been a [c] or [q], with or without aspiration, instead of a [k] (cf. Cincius 1975).

4 Complexification or a change for the worse?

To sum up the discussion thus far, there are arguments for the existence of a phoneme *K- in proto-Tungusic that was lost in the majority of the daughter languages. Its possible existence in a Jurchenic language provides additional evidence against a later innovation (i.e., a prothetic development) and for its potential reconstruction as a plosive. Given that it used to have the function of a submorpheme in the interrogative system, its loss was more than a mere phonological change but also had functional implications.

More specifically, it had the consequence of making the interrogative system more complex on most or all of the seven dimensions mentioned in §2. Arguably, the interrogative system in Nanaic is more regular, less redundant, more analyzable, more organized, more coherent, and better delineated than that of most other Tungusic languages.

First, the interrogative system became *irregular* due to many exceptions from newly created resonances, such as *I-* in Aoluguya Evenki (see §3). Some languages, such as Udihe, lack a resonance entirely, i.e. there is no formal regularity in the first place.

Second, the new interrogative systems are *redundant* in sometimes having more than one resonance, e.g. *a~* and *y~* in Sibe (see below), although the spread of the resonance *y~* might have been independent of the phonological change observed in this paper.

Third, forms that used to be at least partly analyzable (e.g., Nanai *xado*, *xai*) became entirely *unanalyzable* and etymologically opaque (e.g., Manchu *udu*, *ai*).

Fourth, in some languages this loss of analyzability led to an *increase in the number* of interrogatives, especially if the resonance in *K- is considered some form of partially analyzable interrogative stem in its own right. If, on the other hand, the resonance is not granted such a position, its loss did not necessarily affect the number of interrogatives.

Fifth, at a first glance, the overall organization of the interrogatives appears to be unaffected. Because Tungusic languages have suffixes exclusively, inflectional paradigms and derivations generally remained intact. However, from the

point of view of organization, the special position of Tungusic **ŋüi* ‘who’, which is even more pronounced in Nanaic, could also have its merit if this mirrors a special and salient position of the category PERSON in human cognition. In fact, there is empirical evidence for this assumption. In many languages, the personal interrogative stands apart phonologically or morphosyntactically from the rest of the interrogative system (Hölzl 2018b: 406). In addition, few languages have one category for both PERSON and THING and innovative interrogative systems such as in Manchu are usually based on ‘what’ or ‘which’, but rarely on other categories (e.g., Cysouw 2007). An interrogative system as in Uilta with a special position of the category PERSON and a larger set of forms with a shared origin is thus a very organized and natural outcome of general processes and tendencies. It indicates that in pre-proto-Tungusic times there may have been an innovative interrogative system with a large set of analyzable forms that resulted in the later resonance. Given that the resonance covers a historically grown (and ideally coherent) region in the semantic space of interrogatives (e.g., Cysouw 2005, Hölzl 2018b), its loss entirely *disrupted the organization* (i.e., the form-function mapping) of that system (but see below).

Sixth, and most importantly, the formal coherence of the interrogative system was lost. The new interrogative systems simply have no phonological or morphological marker in common but consist of a loose set of synchronically unrelated and *incoherent* forms that only share some semantic similarities.

At a first glance, the seventh dimension appears to be similarly unaffected as the fifth. The phoneme *K-* in Tungusic was not restricted to its function as resonance, but also occurred in many other lexical items (e.g., Ikegami 1997: 227–250). Although the same is true for the new systems, they are much less homogenous and therefore *less delineated* if taken as a whole. For example, there are chance resemblances to the demonstrative systems (e.g., Manchu *uttu* ‘thus, like this’, *udu* ‘how many’) and many lexical items.

The title of this volume is *Language change for the worse*. In the description of the workshop it is based on, “changes for the worse” were defined as those changes “that do not readily follow from an improvement in some other area of the language system”. The complexification of the Tungusic interrogative system is an epiphenomenon of a phonological change. Even if the change in Tungusic had the consequence of a complexification – whether this is a change for the better or the worse is another question –, it was triggered by another change that may have been somehow beneficial (Dixon 2016: 195). By definition, the change in the interrogative system in Tungusic can only be considered a change for the worse if this phonological change was not an improvement in itself. However, the *evaluation* of the phonological change depends on the perspective taken.

Consider, for example, the so-called *preference laws* by Vennemann (1988), e.g.

A syllable head is the more preferred: (a) the closer the number of speech sounds in the head is to one, (b) the greater the Consonantal Strength value of its onset, and (c) the more sharply the Consonantal Strength drops from the onset toward the Consonantal Strength of the following syllable nucleus. (Vennemann 1988: 13f.)

§4 has shown that the onset in Tungusic was most likely a plosive that changed to a fricative and then disappeared in most languages. In other words, there was a loss of the consonantal strength of the onset, a decrease in difference between onset and nucleus, and finally a loss of the head altogether. Notably, some parts of this change must have occurred not once but several times in the different branches and subbranches of Tungusic. From this perspective, both the phonological change and its implications were a change for the worse.⁸

From a different perspective, however, the lenition of the initial consonant can also be conceptualized “as a successive decrease and loss of muscular activity” (Bybee 2007: 950), i.e. a change for the better, because the articulation requires less effort. From this perspective, the changes in the interrogative system cannot be considered a change for the worse. Depending on which of the two perspectives we prefer, the change in Tungusic can be said to be either for the “better” or for the “worse”. This example nicely illustrates that an evaluation is always based on specific purposes and perspectives. Note that this discussion only shifts the evaluation of the development in the interrogatives to another level. The evaluation would also require a cost-benefit analysis that is almost impossible to achieve. Which is more important, the potential benefit of the phonological change or the functional implications in the interrogative system?

The qualitative evaluation of a language is both a problematic and dangerous endeavor (e.g., Lehmann 2006). In the following, this will be illustrated through a criticism of Dixon (2016: 213), who mentions “some of the features [...] which should be ideally present in every language, to ensure that it is an effective vehicle for identification, cooperation, communication, argumentation, and so on.” For example: “An ideal language will have a separate form for each of the standard interrogative words: ‘who’, ‘what’, ‘which’, ‘where’, ‘when’, ‘why’, ‘how’, ‘how much’, and ‘how many’” (Dixon 2016: 227). This appears to be a derivation of the “One-Meaning-One-Form principle” (e.g., Miestamo 2008: 34). However, Dixon’s (2016) argument is highly problematic. (1) The list of interrogatives is

⁸Some languages such as Manchu dialects potentially have an initial glottal stop instead of the resonance, but this problem requires additional research.

rather arbitrary. Dixon (2016) excludes interrogative words such as ‘to do what’ from the list because of their cross-linguistic rarity. Consequently, the other categories must have been chosen on this criterion as well. However, even Dixon’s (2012: 407) more extensive and otherwise very good discussion fails to give any cross-linguistic data on the frequency of these forms. The list would also require a clearly specified threshold of when any given category is included or not. (2) It is by no means clear what “separate” forms, also called “basic question words” (Hengeveld et al. 2012: 46), are, given that the analysis of interrogatives is often not clear-cut (§2). (3) Languages that lack a “separate” form for any of the categories mentioned above can still be an effective means of communication. For instance, there is no “basic question word” for the locative meaning ‘where’ in Tungusic. Nevertheless, all Tungusic languages have means of expressing the notion. Even if any of the categories would be entirely absent from a given language, it would presumably not have been required by the speech community.

Table 13 lists all nine categories mentioned by Dixon (2016), illustrated with some examples from the Jurchenic language Sibe. The Sibe interrogative system is similarly incoherent as that in Evenki. **ŋüi* is preserved as *və* and **ja-* as *ya* and its derivations. Similar to Manchu, the resonance in **K~* was lost, which made interrogatives such as *_afš* (Manchu *_absi*), *_ai* (Manchu *_ai*), or *_ut* (Manchu *_udu*) unanalyzable.

Table 13: Some interrogatives in Sibe (Zikmundová 2013) with Nanai cognates (Ko & Yurn 2011). Not all forms and variants listed.

Category	Sibe	Nanai
who	<i>və</i>	<i>ui</i>
which	<i>ya</i>	-
how much	<i>yask(ə)</i>	-
where	<i>yet</i>	-
how	<i>_afš</i>	<i>xaosi</i> ‘wither’
what	<i>_ai</i>	<i>xai</i>
when	<i>_aitin</i>	-
why	<i>_aⁿ</i>	-
how many	<i>_ut</i>	<i>xado</i>

According to Dixon (2016), the Sibe interrogative system would most likely be considered “ideal” because no form is synchronically analyzable (one form, one meaning). However, why should only these categories be considered and not, say,

‘which one’. In Sibe, this category is expressed with the form *yam(kəⁿ)*, which can be partly analyzed as *ya + əm(kəⁿ)* ‘which + one’. In fact, the form *ya əmkəⁿ* is also attested. Thus, depending on the choice of the categories, the system can be said to be more or less ideal. Most certainly, a language-specific approach that takes into account the whole interrogative system would be more beneficial. For example, Sibe also has additional interrogatives that should be considered (e.g., *ailian* ‘what kind of’).

The loss of the resonance in Tungusic lead to separate forms that are synchronically unrelated. Following Dixon (2016), this should be considered a change for the better. However, as pointed out in §2, *analyzability* is only one of several dimensions. If the dimension of *organization* is taken into account, for instance, the analyzability of certain forms could well be a desirable factor. For instance, Sibe *yet* is still partly analyzable as *ya + DAT* and corresponds to Manchu *ya-de*. Given that this form is marked for case, it is part of a paradigm, e.g. Manchu accusative *ya-be* etc. A decrease in analyzability would certainly make the paradigm less organized, irregular, and thus more complex.

Instead of an evaluation, this chapter tried to *objectively* describe the complexity of the Tungusic interrogatives. Both evaluation and complexity can be applied locally or globally and both are graded categories that can be shown on a scale. However, evaluation is necessarily relative to a certain perspective (Lehmann 2006), while complexity is perhaps best described in absolute terms (Miestamo 2008). Bybee (2015: 10) is certainly correct in her assessment that changes as such “are natural to language and they are neither good nor bad.” A language can only be better or worse for a specific purpose, e.g. expressibility, acquirability, processing, articulation etc. Whether the change in Tungusic is for the better or for the worse can be answered either way, depending on the perspective taken.

5 Conclusion

This chapter is a case study of complexification. It was mostly concerned with a phonological change in Tungusic languages spoken in Northeast Asia and its functional implications in a subsystem of these languages. Several problems were left open and require additional research. It has been shown that the initial was most likely a velar-like plosive, but its exact place and manner of articulation have yet to be identified. It is also an open question what was the original reason for the sound change and whether some sort of language contact may have been involved. Nevertheless, some problems such as the putative reflex *n-* in Manchu could be solved and some new evidence from a relatively unknown Tungusic language called Alchuka was presented.

Acknowledgements

I want to thank Bill Rozycki, the editors, and two anonymous reviewers for reading an earlier version of this paper.

References

- Aixinjueluo Yingsheng. 2014. *Manyu kouyu yindian*. Peking: Huayi chubanshe.
- Alonso de la Fuente, José Andrés. 2011. Venjukov's [1858] 1862/1868 Nanai materials. *Journal de la Société Finno-Ougrienne* 93. 11–35.
- An Jun. 1986. *Hezheyu jianzhi*. Peking: Minzu chubanshe.
- Anderson, Gregory D. S. 1998. *Xakas*. München: Lincom Europa.
- Avrorin, Valentin A. & Boris V. Boldyrev. 2001. *Grammatika orochskogo jazyka*. Novosibirsk: Izdat.
- Benzing, Johannes. 1955. *Lamutische Grammatik: Mit Bibliographie, Sprachproben und Glossar*. Wiesbaden: Harrassowitz.
- Benzing, Johannes. 1956. *Die tungusischen Sprachen: Versuch einer vergleichenden Grammatik*. Mainz: Akademie der Wissenschaften und der Literatur.
- Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 3: *Grammatical categories and the lexicon*, 2nd edn., 169–240. Cambridge: Cambridge University Press.
- Bickerton, Derek. 2016 [1981]. *Roots of language*. Berlin: Language Science Press.
- Bulatova, Nadezhda & Stefano Cotrozzi. 2004. *Ewenkische Erzählungen aus der Insel Sachalin*. Bonn: Verlag für Kultur und Wissenschaft.
- Bybee, Joan. 2007. Diachronic linguistics. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of Cognitive Linguistics*, 945–987. Oxford: Oxford University Press.
- Bybee, Joan. 2015. *Language change*. Cambridge: Cambridge University Press.
- Chaoke, D. O. 2007. *Nanmu elunchunyu yanjiu*. Peking: Minzu chubanshe.
- Cincius, Vera I. 1949. *Sravnitel'naja fonetika tunguso-man'chzhurskikh jazykov*. Leningrad: Nauka.
- Cincius, Vera I. 1975. On the pre-Altaic system of consonants. In Louis Ligeti (ed.), *Researches in Altaic languages*, 299–306. Budapest: Akadémiai Kiadó.
- Cysouw, Michael. 2005. *The typology of content interrogatives*. Paper presented at Association for Linguistic Typology, 21–25 July 2005, Padang.
- Cysouw, Michael. 2007. Content interrogatives in Pichis Ashéninka: Corpus study and typological comparison. *International Journal of American Linguistics* 73(2). 133–163.

- Diessel, Holger. 2003. The relationship between demonstratives and interrogatives. *Studies in Language* 27(3). 635–655.
- Dixon, Robert M. W. 2012. *Basic linguistic theory*, vol. 3: *Further grammatical topics*. Oxford: Oxford University Press.
- Dixon, Robert M. W. 2016. *Are some languages better than others?* Oxford: Oxford University Press.
- Doerfer, Gerhard. 1973. Das Kur-Urmiische und seine Verwandten. *Zentralasiatische Studien* 7. 567–599.
- Doerfer, Gerhard. 1985. *Mongolo-Tungusica* (Tungusica 3). Wiesbaden: Harrassowitz.
- Georg, Stefan. 2004. Unreclassifying Tungusic. In Carsten Naehrer (ed.), *Proceedings of the first international conference on Manchu-Tungus studies (Bonn, August 28 - September 1, 2000)* (Tunguso Sibirica 9), 45–57. vol. 2. Wiesbaden: Harrassowitz.
- Givón, Talmy. 1971. Historical syntax and synchronic morphology: An archaeologist's field trip. *Chicago Linguistic Society* 7. 394–415.
- Grube, Wilhelm. 1896. *Die Sprache und Schrift der Jučen*. Leipzig: Harrassowitz.
- Grube, Wilhelm. 1900. *Goldisch-deutsches Wörterverzeichnis*. Sankt Petersburg: Bei den Commissionaren der Kaiserlichen Akademie der Wissenschaften.
- Hasibate'er. 2016. *Aoluguya fangyan yanjiu*. Peking: Minzu chubanshe.
- Hengeveld, Kees, Maria Luiza Braga, Elisiane de Melo Barbosa, Jaqueline Silveira Coriolano, Juliana Jesuino da Costa, Mariana de Souza Martins, Diego Leite de Oliveira, Vinicius Maciel de Oliveira, Luana Gomes Pereira, Liliane Santana, Cassiano Luiz do Carmo Santos & Viviane dos Ramos Soares. 2012. Semantic categories in the indigenous languages of Brazil. *Functions of Language* 19(1). 38–57.
- Hölzl, Andreas. 2017. New evidence on Para-Mongolic numerals. *Journal de la Société Finno-Ougrienne* 96. 21–37.
- Hölzl, Andreas. 2018. *Das Mandschurische: Ein diachroner Überblick*. Paper presented at Asien-Orient Institut der Universität Zürich, 2018.03.16, Zürich.
- Hölzl, Andreas. 2018b. *A typology of questions in Northeast Asia and beyond: An ecological perspective*. Berlin: Language Science Press.
- Hölzl, Andreas. 2018a. Udi, Udihe and the language(s) of the Kyakala. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 15. 111–146.
- Hölzl, Andreas & Yadi Hölzl. 2019. A wedding ceremony of the Kyakala in China: Language and ritual. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 16. 87–144.
- Hu Zengyi. 2001. *Elunchunyu yanjiu*. Peking: Renmin chubanshe.

- Hu Zhenhua & Guy Imart. 1987. Fu-yü Gïrgis: A tentative description of the easternmost Turkic language. *Papers on Inner Asia* 8. 1–61.
- Ikegami, Jirō. 1974. Versuch einer Klassifikation der tungusischen Sprachen. In Georg Hazai & Peter Zieme (eds.), *Sprache, Geschichte und Kultur der altaischen Völker*, 271–272. Berlin: Akademie Verlag.
- Ikegami, Jirō. 1997. *Uirutago jiten*. Sapporo: Hokkaidō University Press.
- Janhunen, Juha. 1991. *Material on Manchurian Khamnigan Evenki*. Helsinki: University of Helsinki.
- Janhunen, Juha. 2012. The expansion of Tungusic as an ethnic and linguistic process. In Andrej Malchukov & Lindsay J. Whaley (eds.), *Recent advances in Tungusic linguistics*, 5–16. Wiesbaden: Harrassowitz.
- Janhunen, Juha. 2017. Issues of comparative Uralic and Altaic studies (1) : The case of Proto-Mongolic *x. *Hoppō jinbun kenkyū* 10. 97–104.
- Jettmar, H. M. 1937. Der Stamm der oberen Golden (Die Sungari-Ussuri Gruppe). *Mitteilungen der Anthropologischen Gesellschaft in Wien* 67(5/6). 245–276.
- Karlsson, Fred, Matti Miestamo & Kaius Sinnemäki. 2008. Introduction: The problem of language complexity. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, vii–xiv. Amsterdam: Benjamins.
- Kiyose, Gisaburō N. 1977. *A study of the Jurchen language and script: Reconstruction and decipherment*. Kyōto: Shōwado.
- Kiyose, Gisaburō N. 1996. Uvular/post-velar q, ɣ and χ in Jurchen and Manchu. In Giovanni Stary (ed.), *Proceedings of the 38th Permanent International Altaistic Conference (PIAC)*, 237–241. Wiesbaden: Harrassowitz.
- Kiyose, Gisaburō N. 2000. Genealogical relationship of Jurchen dialects and literary Manchu. *Central Asiatic Journal* 44(2). 177–189.
- Ko, Dongho & Gyudong Yurn. 2011. *A description of Najkhin Nanai*. Seoul: Seoul National University Press.
- Langacker, Ronald. 2001. What WH means. In Alan Cienki, Barbara J. Luka & Michael B. Smith (eds.), *Conceptual and discourse factors in linguistic structure*, 137–152. Stanford: CSLI Publications.
- Lehmann, Christian. 2006. The value of a language. *Folia Linguistica* 40(3/4). 207–238.
- Ling Chunsheng. 1934. *Songhuaqiang xiayou de hezhezu*. Nanjing: Guoli zhongyang yanjiuyuan lishi yuyan yanjiusuo.
- Mackenzie, J. Lachlan. 2009. Content interrogatives in a sample of 50 languages. *Lingua* 119. 1131–1163.

- Majewicz, Alfred F. (ed.). 2011. *The collected works of Bronisław Piłsudski, vol 4: Materials for the study of Tungusic languages and folklore*. Berlin: Mouton de Gruyter.
- McWhorter, John. 2007. *Language interrupted: Signs of non-native acquisition in standard language grammars*. Oxford: Oxford University Press.
- Miestamo, Matti. 2008. Grammatical complexity in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, 23–41. Amsterdam: Benjamins.
- Mu Yejun. 1985. Alechuka manyu yuyin jianlun. *Manyu yanjiu* 1. 5–15.
- Mu Yejun. 1986. Alechuka manyu de shuci yu gezhuci. *Manyu yanjiu* 2. 2–17.
- Mu Yejun. 1987. Balayu. *Manyu yanjiu* 2. 2–31, 128.
- Mu Yejun. 1988. Alechukayu yuanyin fasheng de yinbian tedian. *Manyu yanjiu* 7. 1–24.
- Muysken, Pieter & Norval Smith. 1990. Question words in pidgin and creole languages. *Linguistics* 28. 883–903.
- Nakanome, Akira. 1928. *Grammatik der Oroko-Sprache*. Osaka: The Osaka Asiatic Society.
- Nikolaeva, Irina A. & Maria Tolskaya. 2001. *A grammar of Udihe*. Berlin: Mouton de Gruyter.
- Norman, Donald A. 2010. *Living with complexity*. London: The MIT Press.
- Norman, Jerry. 2004. Notes on Manchu pronunciation. *Saksaha: A Review of Manchu Studies* 9(5). 25–30.
- Norman, Jerry. 2013. *A comprehensive Manchu-English dictionary*. Cambridge: Harvard University Asia Center.
- Rozycki, William. 1993. A problem in Proto-Tungus reconstruction: The phoneme *k'-. *Ural-Altaische Jahrbücher* 12. 203–212.
- Schmidt, Peter. 1923. The language of the Olchas. *Acta Universitatis Latviensis* 8. 229–288.
- Schmidt, Peter. 1928. The language of the Samagirs. *Acta Universitatis Latviensis* 19. 219–49.
- Sem, Lidija I. 1976. *Očerki dialektov nanajskogo jazyka: bikinskij (ussuriskij) dialekt*. Leningrad: Nauka.
- Shirokogoroff. 1931. Notes on the bilabialization and aspiration of the vowels in the Tungus languages. *Rocznik Orientalistyczny* 7. 235–263.
- Sunik, Orest. 1958. *Kur-urmijskij dialekt*. Leningrad: Nauka.
- Tang, John. 2011. On the terms for goat recorded in Khitan and Jurchen scripts. *Altai Hakpo* 21. 191–203.
- Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.

- Tsumagari, Toshiro. 2009. Grammatical outline of Uilta (revised). *Journal of the Graduate School of Letters* 4. 1–21.
- Turvey, Michael T. 2009. On the notion and implications of organism-environment system. *Ecological Psychology* 21(2). 97–111.
- Vasilevich, Glafira M. 1958. *Evenkijsko-russkij slovar'*. Moscow: Gosudarstvennoe Izdat. Inostrannykh i Nacionalnykh Slovar'ej.
- Venjukov, M. 1862. Vokabular der Golde-Sprache. *Petermanns Geographische Mitteilungen* 8. 109.
- Vennemann, Theo. 1988. *Preference laws for syllable structure*. Berlin: Mouton de Gruyter.
- Wang Qingfeng. 2005. *Manyu yanjiu*. Peking: Minzu chubanshe.
- Yamakoshi, Yasuhiro. 2007. Hamuniga mongurugo tekisuto: Nichijō kaiwa o daizai ni shita kihon bunrei-shū. In Toshiro Tsumagari (ed.), *Languages of the North Pacific Rim*, vol. 14, 119–157. Sapporo: Graduate School of Letters, Hokkaidō University.
- Yamakoshi, Yasuhiro. 2011. Shineken Buryat. In Yasuhiro Yamakoshi (ed.), *Grammatical sketches from the field*, 137–177. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Zhao Jie. 1989. *Xiandai manyu yanjiu*. Peking: Minzu chubanshe.
- Zikmundová, Veronika. 2013. *Spoken Sibe: Morphology of the inflected parts of speech*. Prague: Karolinum.