

Regional variation in Kurmanji: A preliminary classification of dialects

ERGİN ÖPENGİN*
GEOFFREY HAIG*

Abstract

Investigation of the regional variation in Kurmanji, especially its varieties spoken in Turkey, has been almost entirely neglected in the existing literature on Kurdish. In addition to earlier isolated examinations of Kurmanji dialects (cf. MacKenzie, 1961; Ritter, 1971, 1976; Blau, 1975; Jastrow, 1977), native-speaker researchers have recently provided a substantial amount of dialect material across the Kurmanji-speech zone. However, a methodologically-informed evaluation of these observations into a dialect classification is yet to be undertaken. This article aims at providing an initial classification of Kurmanji-internal variation into major regional dialects, based on lexical, phonological and morphosyntactic data collected from five localities in South-eastern Turkey.

Keywords: Kurdish; Kurmanji; regional variation; dialect; linguistic distance.

Cihêrengiya zimanî ya navxweyî di kurmanciyê de: tesnîfeke seretayî ya zaravayan

Di nav xebatên li ser zimanê kurdî de, heta niha, vekolîna cudatîyên devok û zaravayên kurmanciyê, bi taybetî ewên di nav sinorên Tirkîyeyê de, hema bi temamî hatiye piştguhkirin. Ji bilî çend xebatên serbixwe yên pêştir li ser zaravayên kurmançî (wek MacKenzie 1961; Ritter, 1971 û 1976; Blau, 1976; Jastrow 1977), di nav van salên dawî de vekolêrên kurdîziman qewareyeke mezin a dane û materyelên ji gelek zaravayên kurmanciyê berhev kirine. Lê belê, hêj ev çavdêriyên berbelav bi rengê metodolojîk nehatine nirxandin ku tesnîfeke zaravayan jê bi dest bikeve. Ev meqale dil dike tesnîfeke seretayî ya zaravayên serêkî yên kurmanciyê pêşkêş bike li ser bingehê daneyên peyvi û fonolojîk û rêzimanî yên li pênc deverên başûr-rojhilatê Tirkîyeyê berhevkirî.

جیاوازی ناوچهیی له‌ناو کرمانجیدا: پۆلێنبه‌ندیه‌کی سه‌رتایێ زاراوه‌کان

له‌ناو ئه‌و لیکۆلینه‌وانه‌ی که سه‌بارت به‌ جیاوازی ناوچهیی له‌ کرمانجیدا نه‌جماو دراوه، به‌تایبه‌ت ئه‌وانه‌ی که له‌سه‌ر جۆربه‌جۆری ئه‌و [زاراوه‌ی] که له‌ تورکیا قسه‌یان پێ ده‌کری، به‌ته‌واوی له‌سه‌ر هه‌بوونی ئه‌ده‌بیاتیکی هه‌بوو به‌ زمانێ کوردی چاوپۆشی ده‌کری. سه‌ره‌رای چهند تاقیه‌کردنه‌وه‌یه‌کی تاک و ته‌را سه‌بارت به‌ شه‌وه‌زارمکانی کرمانجی که پێشتر به‌نه‌جام گه‌یشتوون (به‌سه‌روردی بکه‌ن له‌گه‌ڵ مه‌که‌ئینزی، 1961؛ ریتتر، 1971 و 1976؛ بلو، 1976؛ یاسترو، 1977) له‌م سالانه‌ی دوایدا نوێه‌رانی کورد راده‌یه‌کی به‌ه‌ر چاوه‌ له‌ ماده‌ی پێویست سه‌بارت به‌ شه‌وه‌زارمکانی کرمانجیان له‌ ده‌قه‌ری کرمانجی ناخێودا ده‌سته‌به‌ر کردوه. هه‌ر چهند به‌سه‌رسته‌ی پۆلێنبه‌ندی زاراوه‌کان، هه‌تاگوو نه‌یستا، هه‌له‌سه‌نگاندنێکی میتۆدیک و پرزانباری له‌سه‌ر ئه‌م ته‌ئیرامانه‌ نه‌جماو نه‌دراوه. ئامانجی ئه‌م وتاره‌ بریتیه‌ له‌وه‌ی که پۆلێنبه‌ندیه‌کی سه‌رتایێ له‌ جیاوازییه‌ ناوچه‌یه‌ی که له‌ زاراوه‌ سه‌ره‌کیه‌کانی ناوچه‌ کرمانجی ناخێوکه‌ندا هه‌یه‌ به‌ده‌سته‌وه‌ بدات و بۆ ئه‌م مه‌یه‌سته‌ نوێه‌رانه‌ که له‌سه‌ر به‌نه‌مای ئه‌م داتا وشه‌یی، ده‌نگه‌سه‌ی و پێکهاته‌ی ریزمانیه‌یه‌ نه‌جماو ده‌دریت که له‌ پێنج ناوچه‌ی باشووری رۆژه‌لاتی تورکیا کۆ کرانه‌ته‌وه‌.

* Ergin Öpengin is a Postdoctoral researcher and Deutsche Forschungsgemeinschaft fellow at the Department of General Linguistics, University of Bamberg, Germany.
Email: ergin.opengin@uni-bamberg.de.

* Geoffrey Haig is Professor of Linguistics at the Department of General Linguistics, University of Bamberg, Germany. Email: geoffrey.haig@uni-bamberg.de.

Introduction

Like any other natural language, Kurmanji encompasses a considerable spectrum of regional variation. Yet within academia, regional variation in Kurmanji has been almost entirely neglected.¹ There are many reasons for this, not the least the obstructive policies of successive Turkish governments in their refusal to acknowledge Kurdish as a valid object of study – or indeed as a language at all. More generally, much of the Kurmanji speech zone has been characterised by insecurity and violence over the past decades, rendering fieldwork in the region a hazardous undertaking, generally off-limits for mainstream academic funding.

These facts are reflected in the available literature. Neither of the two major available studies of regional variation in Kurdish, MacKenzie's two volumes on Kurdish dialects (MacKenzie, 1961, 1962), and Fattah's (2000) extensive survey of Southern Kurdish, treats Kurmanji in any detail. Fattah deals entirely with Southern Kurdish (Kelhuri, Feyli, etc.), while MacKenzie focussed on Central Kurdish (Sorani); his brief treatment of Northern Kurdish (Kurmanji) was entirely restricted to the Badini dialects of Iraqi Kurdistan. The bulk of the Kurmanji speech zone, covering most of southeastern Turkey and parts of Syria and northwest Iran, has thus remained uncharted territory. More recent overviews (McCarus, 2009) likewise treat Sorani in considerable depth, but have almost nothing to say on Kurmanji. Most recent work on Kurmanji essentially adopts the Bedir-Khan standard, and makes only passing reference to regional variation (e.g. Thackston, 2006). A few isolated studies of individual varieties have been published (cf. Ritter, 1971, 1976 on Midyat; Blau, 1975 on Amadiya and Djabal Sinjar; Jastrow, 1977 on Van), the Kurdish Institute of Paris has produced numerous short sketches, word lists, and collections of proverbs, etc. from local varieties (cf. Enstûtuya Kurdî ya Parîsê, 2010), and an increasing number of native-speaker enthusiasts working on "their" local dialect (e.g. Kömür, 2003) have brought much interesting data to light. But we lack any kind of synthesis for integrating these observations into a larger overall picture, cast in a more consistent linguistic framework.

Our objective with this contribution is thus to provide an initial classification of Kurmanji, based on lexical, phonological and morphosyntactic variation attested in different regional varieties. Given the sheer number of speakers, and the territorial extent of the Kurdish speaking region, the present classification remains of necessity fairly coarse-grained. Nevertheless, the implications of this work go beyond a mere exercise in linguistic taxonomy; the patterns discerned can shed considerable light on the historical processes such as population movements and cultural contacts that have shaped the Kurmanji

¹ We would like to thank an anonymous reviewer, later revealed as Erik Anonby, for his meticulous review, which prompted us to revise the original version in several respects. We would also like to thank Musa Ekici, Musa Aydın, Serdar Ay, Şaziye Şahin for providing us with the data from their dialects. Finally, we thank Nils Schiborr and Maria Vollmer for their assistance with the map and with a number of other formatting matters. The authors alone bear the responsibility for the remaining errors and shortcomings.

speech community over the last 1000 years. While we certainly do not expect this study to be the last word on the internal classification of Kurmanji, we believe that the framework presented here forms a useful provisional classification that will provide a point of departure for future work in this direction.

Our focus is solely on the variation internal to those varieties that are considered to constitute “Kurmanji”. Within the entirety of “Kurdish” (see Haig and Öpengin, this volume, and Jügel, this volume), Kurmanji itself constitutes a relatively well-delineated group; with the exception of some of the south-easternmost varieties of Kurdish such as Surçî (MacKenzie, 1961: 150), analysts have little difficulty in deciding whether a particular variety belongs to Kurmanji or not (see Haig and Öpengin, forthcoming, for a summary of differences between Kurmanji and Central Kurdish). There is in fact considerable variation internal to Kurmanji, but it is no more than would be expected for any natural language spread across a comparable geographic region. Although mutual intelligibility between Kurmanji varieties has not yet been systematically investigated, our impression based on observations in the field over many years is that Kurmanji-internal dialectal variation is seldom a serious obstacle to communication between speakers of different regions.² The dialects of the extreme southeast (our Southeastern Kurmanji) and those of the extreme northwest (our Northwestern Kurmanji) show the greatest divergence from the others, and mutual intelligibility between these peripheral dialects and the others may in fact be restricted; this remains to be investigated.

Today, the existing dialect divisions are progressively blurring due to massive population movements out of Kurdistan, with accompanying language shift and language attrition (see Öpengin, 2012), and also to the emergence of a trans-national urbanised Kurdish culture fuelled by the internet and satellite television. It is therefore a matter of considerable urgency to document as much as possible of the rich regional variety embodied in the dialects before they disappear entirely.

The paper is organised along the following lines. In the second section, we introduce the main geographical divisions that form the basis for our analysis, and outline data sources and means of data compilation. In the third section, we present the results of the comparison of two different kinds of regional variation: lexical variation and phonological differences across cognate forms. In section four, we deal with variation in morphosyntax, while section five summarises the main findings and draws some more general conclusions, both of general methodological nature as well as remarks pertaining specifically to the Kurdish case. The raw data that form the basis of our analysis are compiled in the Appendices.

² As an anonymous reviewer has pointed out, speakers may in fact be resorting to some kind of inter-dialectal, neutral variety when communicating with speakers from other regions, creating the impression that mutual intelligibility is higher than is actually warranted. Whatever the actual mechanisms involved may be, however, we have found little evidence for serious impairment of face-to-face communication due to dialect differences. The same does not hold, however, between Zazaki and Kurmanji.

Methodology and data

A note on transcription

At this point a note on the transcription used is in order: when referring to a particular word in Kurmanji, but outside of any specific local dialect context, we use the standard Roman-based orthography originally developed by Jeladat Ali Bedirkhan at the beginning of 1930s, disseminated in the *Hawar* magazine (cf. Hawar, 1998; Bedir Khan and Lescot, 1991), and widely used in contemporary written Kurmanji, as the neutral pan-regional representation for a word. When discussing actual dialect forms, we deploy the transcription based on the philological tradition of Iranian linguistics (e.g. Mahmudveysi et al., 2012), but we add IPA symbols when necessary for clarification. The transcription used here is broad and generally phonemic; for more phonetic detail on individual regional varieties see Jastrow (1977), Kahn (1976), MacKenzie (1961) and Haig and Öpengin (forthcoming.) The most important conventions employed are the following:

Transcription	Description / IPA symbol
a	open, front, unrounded vowel, may be phonetically long [a, a:]
æ	near-open, front, unrounded vowel [æ]
i	short, central, unrounded vowel, varying degrees of height and backness [i, u, ə]
e	open-mid, front, unrounded, vowel (there is considerable variation in realization) [ɛ]
ē	close-mid, front, unrounded vowel, may be phonetically long [e, e:]
ī	closed, front, unrounded vowel, may be phonetically long [i, i:]
u	short, close-mid, slightly centralised, rounded vowel [ʊ]
ū	close, back, rounded vowel, may be phonetically long [u, u:]
ü	close, front, rounded vowel, may be phonetically long [y, y:]
o	close-mid, back, rounded vowel, may be phonetically long [o, o:]
ɔ	open-mid, back, rounded vowel, may be phonetically long [ɔ:]
superscript h	indicates aspiration of obstruents
h	voiceless pharyngeal fricative [ħ]
ʕ	voiced pharyngeal approximant/fricative [ʕ]
č	voiceless affricate [tʃ]
j	voiced affricate [dʒ]
š	voiceless post-alveolar fricative [ʃ]
ž	voiced post-alveolar fricative [ʒ]
ḫ	voiced uvular fricative [ʁ]
ř	central trill [ɹ]

Main regional divisions

While an empirically grounded dialect classification of Kurmanji does not yet exist, there is a considerable body of perceptions regarding regional variation, and many speakers are familiar with terms such as “Serhed”, “Badini”, or “Botan/Boti”, etc., as designations of regions exhibiting purportedly characteristic linguistic features. To what extent speakers’ perceptions of dialectal

divisions are based on genuine linguistic differences, or merely reflect prevalent cultural/perceived social realities, is a moot question (cf. Preston, 2003), but although the match of speakers' perception to measurable linguistic diversity is not perfect, it is also not entirely random. Therefore, we have initially adopted a division of the Kurmanji speech zone into five regions, based on our own experience with the language over the last decades and reflecting a broad consensus of Kurmanji speakers regarding dialect divisions. These five regions are our points of reference when presenting the data, and we will in fact conclude that they do represent a reasonably well-founded division of Kurmanji into regional dialects. Having decided on the major regions to sample, we proceeded to compile standardised data sets (see below) from a speaker (and sometimes more than one speaker) from each of the five regions. The data sets form the basis for various kinds of analysis, each providing a related, but distinct, measure of linguistic distance.

The five regions identified in our research are the following:

Southeastern dialect region (abbreviated SEK). This region includes the Hakkâri Province of southeastern Turkey and Duhok Province of Iraqi Kurdistan, and includes what is traditionally called the Badini dialect. We have taken the dialect of Şemdinli (Kr.³ *Şemzînan*), a district in the south-east of the Hakkâri Province, as the representative of SEK. The first author of this article, as a native speaker of the dialect, male, 29 years-old, provided the data.

Southern dialect region (SK). This region includes the central-southern section of the Kurmanji speech zone, including the Kurmanji of Mardin and Batman Provinces in Turkey, as well as sections of Şırnak (Kr. *Şîrnex*), some districts of Diyarbakır (Kr. *Diyarbekir*) and Şanlıurfa (Kr. *Riba*) provinces in the Kurdish region in Turkey as well as in Hasaka Province in Syria and the region of Sincar in Iraq. The data for this dialect came from a 28 year-old educated male from the Mardin region.

Northern dialect region (NK). This dialect is commonly referred to as "Serhed" Kurdish, and in Turkey includes the Provinces of Muş (Kr. *Mûş*), Ağrı (Kr. *Agîrî* or *Qerekilîs*), Erzurum (Kr. *Erzerom*) and some districts of the Provinces of Van (Kr. *Wan*), Bitlis (Kr. *Bilîs/Bedlîs*), Bingöl (Kr. *Çewlîg*) and Diyarbakır. The informant is a 40 year-old educated male from Varto (Kr. *Gimgim*), who grew up in the district but has been living outside the language area for the past ten years.

Southwestern dialect region (SWK). This region includes Adıyaman (Kr. *Semsûr*), Gaziantep (Kr. *Entab*) and the western half of Şanlıurfa Provinces of Turkey as well as the northern section of the Aleppo (Kr. *Heleb*) Province in Syria. The data was collected from a 25 year-old educated male from a village in Gaziantep.

Northwestern dialect region (NWK). This region includes the Kurmanji varieties spoken in Kahramanmaraş (Kr. *Meraş*), Malatya (Kr. *Meletî*) and

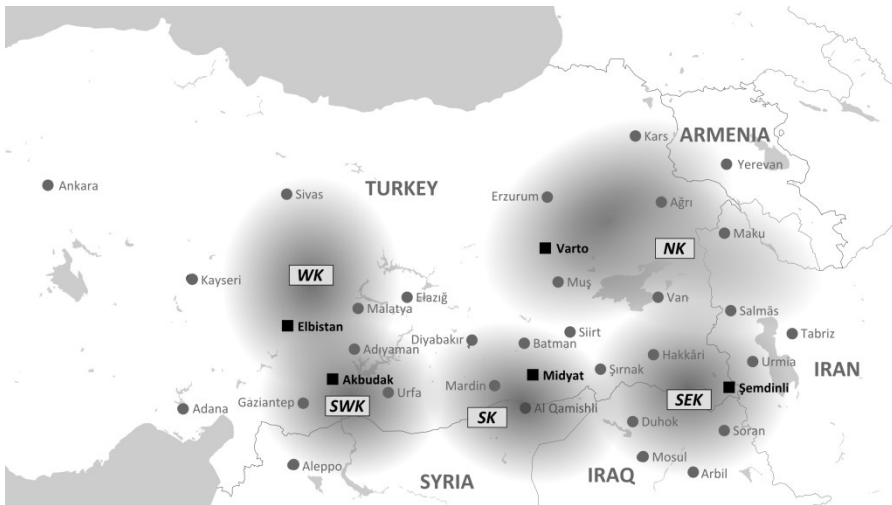
³ Kr. is the abbreviation for Kurdish.

Sivas (Kr. *Sêvaz*) provinces. The data was collected from a 25 year-old educated female from the Elbistan district of Kahramanmaraş.

Standard Kurdish is generally associated with the dialect of Botan region, centred on the town of Cizre (Kr. *Cizîr*), which would be placed between our SK and SEK, and it does in fact show characteristics of both. But although Standard Kurdish is loosely based on the Kurdish of a certain region, it is not a regional dialect on a par with the others discussed here, as its function as the basis for the emergent cross-regional standard (e.g. in print media and satellite TV) confers on it a different status. The dynamics of the development of Standard Kurdish are playing themselves out in a discourse space that cannot be reduced to purely geographic terms, and we therefore refrain from assigning Standard Kurdish to any particular geographic region.

The map in Figure 1 shows the approximate extent of the five regions, and the origins of the speakers from whom the data were collected.

Figure 1. Map of major regional dialects in Kurmanjî



Key: Origins of speakers who provided the data are indicated by squares.

We would like to emphasise the provisional nature of the graphic representation at this stage. First, we are only able to estimate the extents of the respective dialect regions; we currently do not have sufficient data for many regions to make a more precise localisation possible, and we have therefore left some of the Kurdish-speaking regions unassigned to any particular dialect (e.g. west of Diyarbakır). Second, there are regions, e.g. southwest of Lake Van, which appear to lie at the intersection of more than one dialect region; again, resolving the complexities here would require much finer-grained sampling than we have been able to accomplish, and we defer this to later research. Finally, there is a profound problem inherent in this kind of research,

particularly in view of the mobility of the Kurdish community and increasing exposure to other varieties of Kurmanji in the last decades, namely that of variation within a speaker's speech (cf. Kahn, 1976 on variation in Kurmanji phonology). MacKenzie (1961) had already bemoaned the difficulties of finding "pure" dialect speakers. In fact, the ideal of the "pure dialect speaker" is probably an illusion of the analyst, and the precisely localisable isoglosses of traditional dialectology likewise (see Auer, 2005) for a critique of the traditional dialectological notion of isogloss).

Data types

From each region, we elicited four different kinds of data, which are outlined below and available in the Appendices. The data was collected through elicitations with native speakers from the five different regions introduced earlier, gathered with speakers who were living in or near Paris in September–October 2012.

The basic lexicon

It is generally accepted that certain lexical items are less prone to influence from neighbouring languages than others. These words constitute a relatively stable core vocabulary which, all other things being equal, is less likely to be replaced over time than other parts of the lexicon. Typically, this vocabulary includes words for body parts, salient and frequent natural phenomena, verbs for basic activities such as "eat", pronouns, and numerals under 10. A list of 200 such items was originally compiled by the American linguist Morris Swadesh, later reduced to 100 items (the "Swadesh List", cf. Swadesh, 1955). The Swadesh list has been widely applied in a technique known as "glottochronology", a (very controversial) technique in historical linguistics. The underlying assumption is that vocabulary items are replaced at a constant rate across time. Therefore, if we compare the Swadesh vocabulary list in two related languages, we can calculate how many items are shared, and how many differ, which yields a rough indication of the time-span that has elapsed since the two languages split from their common ancestor (Swadesh, 1959). Glottochronology has been severely criticised (cf. Dixon, 1997: 35–36; Campbell, 1998: 177–186; Fox, 1995: 279–291), and few would now take the calculation of absolute dates seriously. In particular, the assumption of a constant rate of vocabulary replacement has been disproved in a number of studies. However, as a method for quantifying "relative" historical distances between related languages or dialects, it retains its value. Furthermore, most linguists concur that some reference to the concept of basic or core vocabulary is relevant for assessing degrees of relatedness, and more sophisticated applications have since been developed (cf. Heggarty, 2012 for an overview). Comparison of the basic lexicon has thus since become one of the tools in the linguists' toolkit for comparing languages, though the original ramifications of glottochronology have largely been abandoned (see Anonby, 2003, 2004/2005 for an application of basic lexicon comparison to Luri and Southern Kurdish).

For our investigation, we used a basic vocabulary list, but rather than use the original Swadesh list, which was based entirely on the intuition of Morris Swadesh, we have adopted the Leipzig-Jakarta list (Haspelmath and Tadmor, 2009). The latter results from a large-scale cross-linguistic investigation of loan words, and is a list of those word-meanings which emerged as the least likely to have been borrowed across the languages in the sample. It is thus an empirically validated/established representation of “basic lexicon”. The list used in this study comprises the 100 items of the Leipzig-Jakarta list, supplemented by a list of words which are specifically relevant to the Iranian languages, yielding 159 in total. The full list can be found in the Appendix A.

Phonological variation

In this list we have included 21 items which are known to show a high degree of phonological variation across different Kurmanji dialects, noting their phonological form in each of the five regions (see Appendix B).

Lexical variation

This list includes 33 items – mostly verbs –, which are characterised by lexical variants across Kurmanji (see Appendix C).

Verb paradigms

This list includes 14 basic verbs conjugated for 2nd and 3rd person singular in indicative present and simple past, as well as in imperative and subjunctive moods (Appendix D).

Analysing variation in lexicon and phonology

Variation can occur in a number of distinct domains. In this study, we focus on three types: (a) variation in the choice of lexical items used to express certain concepts (lexical variants); (b) phonological variation between etymologically related words (cognates)⁴; (c) variation in morphosyntax. Each kind of data has its own advantages and limitations, which is precisely why it is expedient to triangulate from several different methods. If an analysis based on results from one method is replicated by the results of an independent method, then the validity of the original analysis is confirmed to a greater extent than merely further confirmation from the same method. Below we outline the methodology underlying the first two data types (lexicon and phonology), and in section four we take up variation in morphosyntax. In the conclusions we discuss some of the shortcomings, and some avenues for refining and extending the methodology.

⁴ Note that our use of the term “cognate” is a conservative one, reserved for words which are assumed to be reflexes of the same word in the proto language shared by the related languages under investigation. In other studies, “cognate” may also include loanwords which have comparable form and meaning in the related languages under investigation (cf. Gooskens, 2007).

Assessing lexical variation

We use the term “lexical variants” to refer to distinct lexical items expressing the same meaning. With “distinct lexical item”, we mean that to express any particular meaning in the word list, two dialects use words that can be plausibly assumed to have distinct historical origins, i.e. which are not cognates.⁵ This type will be distinguished from phonological differences between two words with the same assumed historical origin (cognates), which are discussed below. It would in fact be possible to collapse the distinction between these two data types (e.g. by working with Levenshtein distances, cf. Gooskens, 2007), but we have opted to preserve the division at this stage. The division will become clearer in our discussion of actual examples.

In order to assess levels of lexical variation, we conducted a pair-wise comparison of each item in the basic vocabulary list from each regional variety with each other variety, and calculated the percentage of the basic lexicon which each pair shared. Thus the percentages in the Table 1 refer to the percentages of the total of 158⁶ items in our basic vocabulary list.

Table 1. Percentages of shared lexical items (cognates) in the 158-item word list across the five regions

	SEK	SK	NK	SWK
SK	79			
NK	73	80		
SWK	73	80	87	
NWK	72	77	81	87

151

The percentages of shared cognates given figures in Table 1 are actually lower than what one could expect for supposedly mutually intelligible dialects. For example, Girdenis and Mažiulis (1994: 9), working with the Swadesh list on Slavic languages, show that Bulgarian and Macedonian share 86% and Polish and Russian 77% of their basic vocabulary. Similarly, Dyen et al. (1992) show that shared basic vocabulary between Spanish and Italian on one hand and French and Italian are at 80% and 79% respectively. A naïve comparison of these figures with ours for Kurmanji would suggest, for example, that the difference between SK and NK (80%, cf. Table 1) is comparable to that between Spanish and Italian.⁷ However, we urge caution in comparing the fig-

⁵ A reviewer suggests that the two kinds of data we are distinguishing here both involve “lexical differences”. We see this basically as a disagreement in terminology; the conceptual difference between cognates and non-cognates, which is behind our distinction, is sufficiently uncontroversial, regardless of the labels chosen to refer to it.

⁶ One item was removed from the original list due to difficulties of analysis.

⁷ A reviewer actually concludes from the figures in Table 1 that the “Kurmanji dialects themselves are not mutually intelligible, but that speakers have learned how to communicate by using interdialects”. This claim is based on the questionable assumption that there is some identifiable threshold of lexical differences that defines “mutual intelligibility”. In fact, existing stud-

ures obtained in different investigations. First, our investigation is not based on the Swadesh list, but on a modified version of the Leipzig/Jakarta list, rendering direct comparison invalid. Second, we have been extremely conservative in considering two words to have the “same” meaning. Thus a relatively slight semantic shift between cognate words in two dialects (e.g. “nose” > “nostril”) is sufficient for us to treat those two dialects as having distinct words for that item on the list, thus lowering the overall percentage of shared items. Other differences in analytic procedure also impact on the results and it is often not possible to see exactly which decisions underly the figures in published sources. What Table 1 provides, then, is not an absolute measure of distance, which would be directly comparable to other investigations, but a measure of relative historical distance among the varieties considered in this investigation.

Turning now to the figures themselves, the “closest” dialect pairs are Northwestern Kurmanji (NWK) and Southwestern Kurmanji (SWK), which share 87% of items in the word list, and Northern Kurmanji and Southwestern Kurmanji (SWK), likewise with 87% shared common lexicon. The comparison of shared lexical items in fact already yields a broad sub-classification into three main dialect groups: the three western and northern varieties NK, NWK and SWK cluster together, in distinction to Southeastern Kurmanji (SEK). Intermediate between these two is Southern Kurmanji (SK), with between 77 and 80% shared similarities to all other varieties.

Table 2 gives a selection of the data, showing examples of variation and commonalities across the five dialects. The lexical items that set one dialect apart from the others are in *italics*. It is evident that different dialect combinations emerge, depending on the lexical item. For instance the words *ber* “stone” and *phir* “hair” set SEK apart from all other dialects; the word *baş* “good” bundles SEK and SK together as different from the other three dialects, which share as equivalent the word *rind*; the word *mezîn* “big” bundles the four dialects together as different from NWK which has *gir* as the corresponding word. Mapping these different forms would not, therefore, yield congruent isoglosses. However, in general we find NWK and SEK pattern as the ends of a southeast/northwest continuum, showing distinct lexical items, while the intermediate dialects align for some items with SEK, for others with NWK, and in a few cases exhibit unique forms (e.g. NK *şîlî* “rain”).

ies on mutual intelligibility conclude that the correlation between lexical differences and intelligibility is “not significant” (Gooskens, 2007: 461); a better match is obtained through phonetic distance based on Levenshtein distances; this is obviously a promising avenue for further research. However, other factors also impact on mutual intelligibility, and actually testing them turns out to be an exceedingly complex task. It is therefore not clear how the reviewer is interpreting the range of values in Table 1 (72-87%) in terms of mutual intelligibility. Given the rather obscure nature of the construct “mutual intelligibility”, we prefer not to interpret the figures in Table 1 in this manner, but simply to take them as an initial – and very approximate – measure of relative historical distance.

Table 2. Examples of lexical variants in Kurmanji dialects

Stand.K + Gloss	SEK	SK	NK	SWK	NWK
<i>kevîr</i> ‘stone’	<i>ber</i>	kevîr	kevîr	kævîr ⁸	kævîr
<i>por</i> ‘hair’	<i>p^hirê</i>	p ^h ov	p ^h or	p ^h or	p ^h or
<i>lêxistin</i> ‘hit’	<i>lêdan</i>	lêxistin	lêxistin	lêxistin	lêxistin
<i>zarok</i> ‘child’	<i>biçûk</i>	zařok	zar	zar	<i>dêl</i>
<i>poz</i> ‘nose’	<i>difîn</i>	<i>behvîl</i>	poz	poz	poz
<i>baş</i> ‘good’	baş	baş	<i>rînd</i>	<i>rînd</i>	<i>rînd</i>
<i>ling</i> ‘leg’	ling	ling	<i>çip</i>	nig	<i>çûnî</i>
<i>baran</i> ‘rain’	baran	baran	<i>şîlî</i>	baran	baran
<i>per</i> ‘wing’	p ^h eř	p ^h eř	p ^h eř	<i>pîl</i>	<i>pîl</i>
<i>mezîn</i> ‘big’	mezîn	mezîn	mezîn	mezîn	<i>gir</i>

The attested differences in the basic lexicon arise primarily through two main mechanisms. The most important is semantic shift: the cognates undergo different changes in meaning in the respective varieties (or a change in one variety, while in another the meaning remains unchanged), leading over time to distinct meanings. As an example, consider the word for “nose” in SEK, *difîn*. In NK, the word for “nose” is *poz* (cf. Table 2). However, the word *difîn* does in fact exist in NK (and probably in other dialects), but it has a more specialised meaning than just “nose”, namely “nostril”, hence it is not included in this list. Further examples of semantic shifts include the words for “stone”: in all dialects it is *kevîr* except for SEK, where we find *ber*. In SEK a cognate of *kevîr* does exist, but it has undergone a semantic shift to “steep cliff”. Similarly, the word for “hair” is *p^hirê* in SEK but *p^hor* in all other dialects. The word *p^hor* exists in SEK with the more specific meaning of “lock of hair above the forehead” while, in turn, the word *p^hirê* exists in the other dialects with specified meaning of “the hair on the body of human beings”.

Many similar examples can be found outside of the basic vocabulary (cf. Appendix C). For example, the verb *şelimandin* exists in SK and NK with the meaning “learn/teach”, but in SEK, it refers to “bringing up (a child)”. The word *peş* is used as the lexical variant of *mî* “sheep” in NWK, but in all other dialects *peş* is the generic term for all types of sheep. Similarly, the word *kûçîk* “dog” in NK and SWK, and its – probable – cognate *kudîk* in NWK, are used more specifically to refer to a “young dog, whelp” in NWK and SEK. The word *biçûk* “child” in SEK is found in all the other dialects, but with the conceptually-related meaning of “small”. Finally, the word *gir* exists in all of the

⁸ We use the symbol [æ] to indicate a more open realization of Standard Kurdish [e], which is rendered in Standard Kurdish orthography as <e> (as in Standard Kurdish *ez* “I”). If a dialectal realisation of this vowel is close to the Standard Kurdish one, we transcribe it with <e>. However, the two-way distinction in our transcription is certainly an over-simplification.

five dialects, but only in NWK is it the general-purpose adjective meaning “big”, while in other dialects it bears a semantic nuance akin to “bulky”.

One might conjecture that variation caused by semantic shifts would be less detrimental to mutual intelligibility because the relevant words are available to the other speakers, and the semantic differences are often fairly transparent. Semantic shifts affecting cognate words are of course characteristic of related varieties, whether dialects or languages. With increasing time of separation of related varieties, semantic change may cause cognates to diverge considerably in their meanings. Consider the Germanic languages English and German; in English we find the word *deer*, which has the cognate in German *Tier* “animal”. Likewise English *thatch* has the German cognate *Dach* “roof”, and English *fool* has as cognate in German *Vogel* “bird”.

The second major source of lexical differences is borrowing from neighbouring languages, mainly from Arabic and Turkish but sometimes also from Armenian, Persian and Sorani (though it is often difficult to prove that an item has been “borrowed” from the latter into Kurmanji). However, most of the clear instances of borrowing occur outside the basic lexicon, as we have defined it above. One of the reasons that SEK appears to diverge most from the other Kurmanji dialects is that it shares many words from neighbouring Sorani (e.g. *ber* “stone”, *bežir* “lost”, *sotin* “burn”, *šimî* “melon”). Whether these are actually “borrowings” or shared retentions that have been retained due to the relative proximity with Sorani, can probably not be established with any certainty. It should be noted that they are found throughout the SEK region (areas like Çatakin, Van), so are certainly not the result of recent direct contact with Sorani.

In SK, we find higher proportion of loans from Arabic, including common verbs such as *šitexlandin* “speak”, *betlandin* “get tired”, *selimandin* “learn”, or *şuxulandin* “work”, fairly obviously triggered by the long-standing coexistence with Arabic-speaking groups in the region. In NWK, Arabic influence is less obvious, while Turkish influence is more apparent, for example expressions like *belli kirin* “learn” based on a Turkish word *belli* “obvious”, or the widespread use of Turkish verb-forms with *-miş* coupled with Kurmanji light verbs, amply attested in Le Coq’s texts published in 1903 (Le Coq, 1903). The different words for “potato” also reflect different sources of borrowing: *kartol* in NK is from a Slavic language, while *patetîz* in SWK is presumably via Turkish. However, while language contact has contributed to lexical variation in Kurmanji, it is noteworthy that few items on our list of basic vocabulary any of the dialects are clearly identifiable as borrowings (borrowings include *sert* “hard”, *qûm* “sand” and *qalçe* “thigh”, from Turkish, and *sekinîn* “to stop” *şo(ş)/* “work” from Arabic). Borrowing has thus – in confirmation of the basic assumption behind the Leipzig-Jakarta List – primarily affected words outside the basic vocabulary.

The most divergent dialect with regard to lexical variation is SEK, which shares less than 80% of its basic lexicon with the other dialects. We should note that our SEK includes what is commonly referred to as Badini, the

Kurmanji of Iraqi Kurdistan. As such, it exhibits many lexical items not found elsewhere in Kurmanji, such as *axiftin* “say”, *şîyan* “be able”, *sêwik* “potato”, *şeqî bûn* “be tired”, *bezîr/berze* “lost”. These items are often considered the hallmark of Badini, setting it off from the Kurmanji spoken in Turkey. However, we would like to emphasise that there is no sharp dialect border corresponding to the national border Turkey/Iraq, and that the Kurmanji of Southeastern Turkey and in the area across the Iranian border shares most of the lexical and grammatical features of the neighbouring varieties of North Iraq. We therefore refer to a broader unit “SEK”, which includes both Badini and the neighbouring Kurmanji varieties of southeastern Turkey and across the border in Iran (cf. Fig. 1 above). In this context it is all the more remarkable that some items in NWK, such as *şûni* “knee” and *wêst* “now”, are shared with Sorani, and *şæ kirin* “can, be able” is reminiscent of the archaic SEK form *şîyan*. Whether these words are hints of an old-layer of migration from further southeast, or retentions that happen to be shared by these two varieties, cannot be readily decided on the basis of the little data available.

In sum, the instances of variation observed in the comparison of the basic lexicon are typical examples of natural processes of language change: semantic shifts, and borrowing. Given sufficient time depth and continued separation, the unrelenting effects of such changes accumulate, and ultimately lead to the emergence of distinct languages. There is, however, no generally accepted answer to the question of at what level such differences should be considered to yield different “languages”. However, the concept of “language” involves not only matters of lexical and morphosyntactic similarity, but also meta-linguistic issues of perceived unity and shared cultural heritage, as discussed in Haig and Öpengin (this issue), and it would be premature to engage in that discussion here. Our findings regarding degree of lexical variation provide an initial, and fairly rough, gauge of historical distance among the dialects investigated. The picture that emerges does in fact reflect the areal distribution: the geographically most peripheral dialects, SEK and NWK, are also those that share the least cognates (72%), which can be read as indicating greatest time-depth of separation. These findings are largely confirmed by other data types considered below.

Phonological variation in cognate words

Regular sound correspondences

Cognates are words in related languages and dialects that are considered to have been inherited from the same word in the common ancestor language (cf. Campbell and Mixco, 2007: 33). In this section, we investigate a selection of cognate words, both within and outside the basic vocabulary, which exhibit systematic phonological variation (Appendix B). Issues of semantic shifts are ignored in this section. Most of the systematic phonological variation found in these words concerns the vowel segments, a finding which confirms Prokić (2010: 142) conclusions from her investigation of Bulgarian dialects. This is of course precisely what one would expect in a comparison of closely

related varieties/dialects: the main locus of accent differences is in vowel qualities, i.e. these are the first phonological segments to diverge over time, with unstressed vowels the least stable of all. Consonants, on the other hand, are relatively stable across related dialects, as Prokić (2010) also confirms. Again in keeping with what is known about comparisons of closely-related varieties, we find that there are relatively few stable and consistent sound correspondences (Prokić and Cysouw, 2013), and most of them concern vowels. Table 3 provides a selection of the correspondences that can be observed.

Table 3. Regular sound correspondences in Kurmanji dialects

Stand.K – example			SEK	SK	NK	SWK	NWK
<i>a</i> [a:]	<i>agir</i>	‘fire’	[a:]	[a:]	[a:]	[a:]	[ɔ:]
<i>e</i> [ɛ]/[æ]	<i>dev</i>	‘mouth’	[ɛ]	[æ] ¹⁰	[ɛ]	[ɛ]	[a:]
<i>û</i> [u:]	<i>gûz</i>	‘walnut’	[y:]	[u:]	[u:]	[u:]	[u:]
<i>o</i> [o:]	<i>iro</i>	‘today’	[u:]	[o:]	[o:]	[o:]	[o:]
<i>VbV</i> [-b-]	<i>bebu⁹</i>	‘there was’	[-b-]	[-b-]	[-b-/-v-]	[-w-]	[-w-]
<i>v</i>	<i>av</i>	‘water’	w	v	v	v	v ¹¹
<i>xw</i>	<i>xwê</i>	‘salt’	x	xw	xw	xw	xw/x ^w

The Stand.K vowel *a*, phonetically [a:], is realised distinctively in NWK as a mid-low back rounded vowel [ɔ:]. Özsoy and Türkylmaz (2006: 304) suggest Turkish influence behind this change, although a change of this nature could be expected within vowel systems without obvious foreign influence, and it is not readily apparent why Turkish influence should be assumed. In accordance with the principles of chain shifts in vowel systems (Labov, 1994), this change is accompanied by a lowering of *e* [ɛ] to *a* [a:] in NWK. Accordingly, the items seen in Stand.K and other dialects as *agir* ‘fire’ and *kevir* ‘stone’ are seen as *ɔgir* and *kavir*. This latter change, *e* > *a* (or *æ*), is also seen in SK but only word-finally and rarely elsewhere (see fn. 10).

The Stand.K *û*, phonetically [u:], is fronted into *î*, phonetically [y:], in SEK. Examples contrasting NK and SEK forms are the following:¹²

⁹ A reviewer questions the validity of this example due the morpheme boundary between *be-* and *-bu*. We acknowledge it as a possible special case, but we retain it in the the table because it is a high-frequency item for which we have reliable data for all dialects. Other words with intervocalic *-b-* generally follow the pattern indicated (see examples further below), though as in most cases of phonological change between closely-related varieties, there may be lexical exceptions and register-determined variation (cf. for example the distribution of diphthong [Au] and monophthong [u:] in words like *down* in Scottish English, described in Smith et al., 2009).

¹⁰ Note that this change in SK affects mostly word-final vowels and sporadically word-initial (*axman* ‘sky’) and word-medial (*mazîn* ‘big’) vowels too. The relevant environment is quite possibly the open syllable (hence it does not effect the SK pronunciation of Stand. K. *dev* ‘mouth’), but this requires more detailed investigation.

¹¹ Unlike in much of West Iranian, [v] and [w] are distinct phonemes in most of Kurmanji, cf. Jastrow (1977), MacKenzie (1961: 30–44) and Haig and Öpengin, forthcoming.

Stand.K		SEK	NK
<i>dūr</i>	‘far’	<i>dūr</i>	<i>Dūr</i>
<i>čūčik</i>	‘bird’	<i>čūčik</i>	<i>čūčik</i>
<i>gūž</i>	‘walnut’	<i>gūž</i>	<i>gūž</i>

This change is followed by raising the close mid *o* to *ū*, so *řož* ‘sun’ of NK is realized as *řūž* in SEK. In the western half of the Badini dialect zone (e.g. Duhok), the change *ū* > *ü* has gone further by derounding the vowel to *i*, e.g. ‘far’ *dūr* > *dūr* > *dür*, or third singular past of ‘be’, which is *bī*. Derounding also affects some items of Şemzinan-dialect of SEK (e.g. *tīž* ‘bitter’, *xîn* ‘blood’). In fact it sporadically affects items further to the west, for example in SK, the word for ‘turtle’ *k̂iso* (cf. Table 4). The fronting (and in some cases derounding) of *ū* is thus a much more complex process than can be adequately treated here, and affects different items to different degrees, but it seems to be more or less restricted to SEK.

SEK does not have the voiced labiodental fricative phoneme /v/ of Stand.K, which has merged with the approximant /w/. Furthermore, the consonant group [xw] is simplified and delabialised to [x] in SEK. Note that these two features distinguish also Central Kurdish (Sorani) from Kurmanji, showing the intermediary position of SEK between Sorani and Kurmanji.

The lenition of Stand. K /b/ to [w] intervocally is one of the most salient features of NWK and SWK. In NK, it is also characteristic for /b/ to weaken to [v] intervocally. Thus it seems that we are dealing with a general process of b-lenition that affects NK, SWK and NWK, and which has proceeded farthest in the latter two dialects (b>v>w). Examples of lexical items from our data are shown below:

Stand.K	SEK	SK	NK	SWK	NWK
hebek ‘one piece’	<i>hebek</i>	<i>hebek</i>	<i>hevek</i>	<i>hewek</i>	<i>hawek</i>
zebeš ‘melon’	-	<i>zebeš</i>	<i>zeveš</i>	<i>zeweš</i>	-

The lenition of [b] is frequently observed with the present-tense stems of verbs beginning with [b-], when they are preceded by the indicative prefix *dī-*, or the subjunctive prefix *bī-*. In this environment, the stem-initial [b-] may weaken further to become a front rounded vowel [œ], for example [dæ:m] ‘I say’, from Stand. K. *dī-bêjim* via lenition of the intervocalic [-b-]>[-w-]>[œ]

¹² The Stand.K. word *xwîn* ‘blood’ appears to have been affected by two of these changes, namely the fronting of the vowel [u:], and the delabialisation of the consonant cluster [xw-], which appear to have interacted in interesting ways. The forms of this word in the various dialects (cf. Table 4) cannot in fact be predicted by any of the rules given in Table 3.

(field notes from Karakoçan). An almost identical process affects initial [b-] of present-stem verbs in the Gorani dialect investigated by Mahmoudveysi et al. (2012: 31).

Other changes in phonological form

Different phonological shapes of cognates may also be motivated by changes other than regular phonological rules: dialects may develop distinct mechanisms for accommodating syllable complexity, or apply different affixes under different conditions, leading to changes in the forms of related words, or there may be sporadic shifts in the phonologies of isolated words (metathesis or dissimilation, for example). Some examples of this kind of variation are illustrated in Table 4.

Table 4. Phonological differences across cognate words in Kurmanji dialects

Stand.K + Gloss	SEK	SK	NK	SWK	NWK
<i>ziman</i> ‘tongue’	<i>ʕezman</i>	<i>ziman</i>	<i>ziman</i>	<i>ziman</i>	<i>ziman</i>
<i>xwîn</i> ‘blood’	<i>xîn</i>	<i>xwîn</i>	<i>xûn</i>	<i>xûn</i>	<i>xûn</i>
<i>bestî</i> ‘bone’	<i>bêstik</i>	<i>bêstû</i>	<i>hestî</i>	<i>hesti</i>	<i>hêstî</i>
<i>spî</i> ‘louse’	<i>hispi</i>	<i>speb</i>	<i>sipî</i>	<i>spî</i>	<i>ispî</i>
<i>zarok</i> ‘child’	<i>biçûk</i>	<i>zařok</i>	<i>zarû</i>	<i>zar</i>	- (dêl)
<i>mîro</i> ‘ant’	<i>mêrû</i>	<i>mîro</i>	<i>mirjolek</i>	<i>mori</i>	- (gêra)
<i>xwîşk</i> ‘sister’	<i>xûşk</i>	<i>xweb</i>	<i>xweng</i>	<i>x^waying</i>	<i>x^wong</i>
<i>diran</i> ‘tooth’	<i>didan</i>	<i>dinan</i>	<i>diran</i>	<i>diran</i>	<i>didon</i>
<i>heyv</i> ‘moon’	<i>bewî</i>	<i>heyv</i>	<i>hîv</i>	<i>bêv</i>	<i>hîv</i>
<i>doh</i> ‘yesterday’	<i>dubu</i>	<i>doh</i>	<i>do</i>	<i>dibu</i>	<i>do</i>
<i>řî</i> ‘beard’	<i>řidin</i>	<i>řih</i>	<i>řû</i>	<i>řibi</i>	<i>řû</i>
<i>keûsi</i> ‘turtle’	<i>keûset</i>	<i>keûso</i>	<i>keûsî</i>	- (řeq)	<i>keûsê</i>

An important source of differences among nouns is the suffix *-(i)k*, sometimes (misleadingly) referred to as the diminutive suffix.¹³ Thus we often find nouns ending in *-k* in one dialect, while in another dialect the cognate noun lacks the *-k*; cf. SEK *bêstik* ‘bone’, or SK *zarok* ‘child’.

In some words, we find reflexes of historical changes that we are unable to systematise. The word for ‘tooth’ had the forms *dnd’n* or *dandan* in Parthian. It shows interesting developments of the medial consonant cluster. In SEK and NWK it is simplified to *-d-*, in SK it is *-n-*, while NK, SWK and Stand.K have *-r-*. There seems little point in postulating directionality of change here; rather, we seem to have parallel processes of simplification in the different dialects. Similarly, the final consonant cluster *-rg-* of the word for ‘wolf’ is simplified in all dialects (as *gur*) except for SEK (*gurg*).

¹³ The function of this Kurmanji suffix is difficult to circumscribe, and its origin is also a puzzle; it may in fact be the reflex of more than one ancient suffix.

Another source of differences concerns certain word-initial consonant clusters, as shown in the dialect forms of the Standard Kurdish *ziman* ‘tongue’, *şpi* ‘louse’ and *stêr* ‘star’. No dialect tolerates an initial [zm-] cluster, but different strategies are employed for handling it: in SEK an initial vowel is added (i.e. *izman* or *ezman*), and the resulting syllable also receives an onset through the insertion of a glottal stop (which may be realised as a pharyngeal fricative [ʕ]), yielding *ʕezman*. SEK also favours this strategy for [sp-], though the onset is provided by [h] rather than a glottal stop. In the other dialects, [zm-] is broken up by an epenthetic vowel (*ziman* etc.). The strategies for dealing with the initial cluster in *şpi*, on the other hand, are quite varied, and both SK and SWK in fact tolerate the cluster in this word. Finally, the cluster *st-* (as in the word for ‘star’) is retained in SEK and SK and avoided in other dialects by inserting a central vowel. Thus, a given dialect may not follow a systematic strategy of handling the initial consonant clusters.

The pharyngeal consonant [h] is found in a number of native Kurmanji words, especially in initial position, for example [hæft] ‘seven’ in SEK. Similarly, loan words from Arabic may retain their pharyngeal consonants. However, in NK we find that pharyngealisation may also affect the quality of unrounded vowels (or of the entire syllable) of a number of words, mostly of native origin, presented in Table 5.¹⁴ A similar phenomenon also affects some words in SK and SWK to differing degrees. It is altogether absent in SEK and NWK, though in the latter, similar to the dialects with pharyngealised vowels, the initial stops are deaspirated. It is typical for SEK that deaspiration of stops and affricates affects fewer lexical items, and is perceptually less salient than in the other dialects. Again, this seems to reflect the proximity of SEK to Sorani Kurdish, which lacks the aspirated/non-aspirated distinction on stops and affricates completely.

Table 5. Pharyngealisation in Kurmanji dialects

Stand.K		SEK	SK	NK	SWK	NWK
<i>çav</i>	‘eye’	[tʃ ^h a:v]	[tʃa:ʕv]	[tʃeʕv]	[tʃa:v]	[tʃɔ:v]
<i>pehn</i>	‘wide’	[p ^h a:n]	[pehn]	[peʕn]	[peʕn]	[pɔ:n]
<i>tehl</i>	‘bitter’	[t ^h a:l]	[tehl]	[teʕl]	[teʕl]	[tɔ:l]
<i>masî</i>	‘fish’	[ma:si:]	[ma:si:]	[meʕsi:]	[ma:si]	[mɔ:si:]
<i>mar</i>	‘snake’	[ma:r]	[ma:r]	[me:ʕr]	[meʕr]	[mɔ:r]

Finally, we should mention a lexico-grammatical feature that distinguishes between different dialects, namely the form of the demonstrative particle, as

¹⁴ In some cases it can be linked to the ‘ejective’ character (lack of aspiration) of the initial consonant, which appears to be re-interpreted as a pharyngeal characteristic, and then spread across the entire syllable. However, other words with a pharyngeal vowel quality such as *mar* ‘snake’, or *mehin* ‘mare’ (from NK) lack an original unaspirated initial consonant, so we lack an explanation for the source of pharyngealisation here.

lexicalised in some adverbs. As shown in Table 6, the demonstrative *ew* ‘this’ is lexicalised as *î-* in SK and NK, along with Stand.K., while the SWK and NWK have it as *hi-* and SEK in its less-modified form *ew* which it shares with Sorani.

Table 6. Demonstrative particle in adverbs

Stand.K		SEK	SK	NK	SWK	NWK
<i>îşev</i>	‘tonight’	ewşew	îşev	îşev	hişev	hişev
<i>îro</i>	‘today’	ewrô	îro	îro	hiro	hiro

Summary of variation in lexicon and phonology

The preceding sections have identified a number of domains of regional variation, focussing on lexical variants and phonological variation in related words (cognates). As mentioned earlier, our initial classification into regional dialects, which formed the basis for data collection, is based on shared folk perceptions and our own knowledge. However, our impression after evaluating the data is that on the whole, the five regions initially identified can in fact be maintained, as they show sufficient internal commonalities, and sufficient differences to the others to render the classification meaningful. While it is evident that the basic lexicon shows a high degree of shared items, outside the basic lexicon the dialects investigated here have a fair number of differences, some of which are captured in the data compiled in Appendices B and C.

Table 7. Lexical isoglosses (lexical variants and cognates) in Kurmanji dialects

Item	SEK	SK	NK	SWK	NWK
‘stone’	ber	kevir			
‘lost’	bezir	winda			
‘much’	hind	qas			
‘hungry’	birsî	birçî			
‘burn’	sotin	şewitandin			
‘arm’	Mil		Pî		
‘like’	Wek		Mina		
‘all’	Hemi		giş(t)		
‘fire’	Agir			Ar	
‘today’	ewro/îro			Hiro	
‘eleven’	Yanzdeh			de-w-yêk	
‘leg’	Ling				žuni
‘now’	Niha				wêsta
‘fly’	Mêş				meş

Table 7 gives a summary of some salient differences, showing that within our dialect divisions different lexical isoglosses bundle the dialects in different

ways. What we essentially find – and this is again as expected – is clear evidence of a dialect continuum, with the intermediate dialects SK, NK and SWK, exhibiting greater commonalities with other dialects than the peripheral varieties SEK and NWK. We nevertheless find evidence for distinguishing NK from SK, and NK from SWK, thus confirming the original choice of dialect divisions.

Variation in morphosyntax

Due to time restrictions, we were unable to compile comprehensive data from all five dialect regions for morphosyntax. In this section, we will illustrate grammatical variation for only three of the main dialects, namely SEK, SK and NWK, on an east-to-west line. NK and SWK, on the other hand, generally exhibit a combination of the variants identified for the three main dialects, which is consonant with the intermediate position of the NK and SWK dialects as outlined above. In this section we treat only a selection of the actually attested variation in morphosyntax; the reader is referred to Haig and Öpengin (forthcoming) for more extensive discussion and illustration of these and additional features.

Personal and reflexive pronouns

The direct/oblique distinction in all pronouns is retained in all dialects.¹⁵ Substantial formal differences are seen mainly in 2nd person plural. As shown below, the SEK form (shared also in northern varieties of Sorani Kurdish) is remarkably different, especially in its oblique form, from the other dialects – although it is possible that they all derive from the same etymological source.

Stand.K	SEK	SK	NWK
<i>bun/we</i> ‘you.pl.dir/obl’	hing/hingo	win/we	wun/we

There is an invariant reflexive pronoun in Kurmanji, used also as an adnominal possessor. In Stand.K and most dialects of Kurmanji, the reflexive pronoun is used as an adnominal possessor only under the condition that its coreferent antecedent is the subject of the same clause (i.e. binding principles). In NWK, however, the reflexive pronoun is generalised as a marker of a third person (singular/plural) possessor, regardless of whether it is coreferential with the same-clause subject. Thus, a clause like *birayê xwe rind-e* (brother-of self/reflexive good-is) ‘his brother is good’ is possible in NWK, but impossible in SEK or SK (which could only have *birayê wî baş-e* for the same meaning). This is in fact a highly salient syntactic difference, as the rule for coreferential binding of reflexive *xwe* is otherwise a very robust feature defining Kurmanji syntax (Haig, 1998).

¹⁵ Though Dorleijn (1996) reports syncretism in the second person singular between Stand.K. *tu* (direct) and *te* (oblique) in the Diyarbakir variety.

Verbal agreement in the present tense

Verbs in Kurmanji inflect for person and number of the subject in present tenses. However, there are a number of regional differences in the respective paradigms. Table 8 presents the main areas of variation, disregarding the variation caused by regular sound changes and certain other details.

Table 8. Verbal person marker paradigms in Kurmanji dialects

	1sg	2sg	3sg	1pl	2pl/3pl
Stand.K	-im	-î	-e	-in	-in
SEK	-im	-î	-î t	-î n	-in
SK	-im	- e	- ê	-in	-in
NWK	-im	- e	- i	-in	-in

The comparison of the data shows that SEK is the most divergent of the dialects in this respect: SEK has an additional number distinction in plural persons (1pl *-î**n***) while the other dialects have a shared plural ending *-in* for all persons. Similarly, the retention of the final consonant of the third person singular ending is a salient dialectal feature setting SEK apart from all other dialects.¹⁶

Case marking and plural ezafe

All dialects indicate oblique case on singular feminine and plural nouns, and all indefinites, but show differences in the way masculine singular nouns are marked for case. Here we only mention two possibilities: oblique case via a suffix *-î* (*nan*, *nan-î*) ‘bread, food’ (direct and oblique), or via raising of the stem vowel (*nan*, *nên*); see Haig and Öpengin (forthcoming) for details. In the ezafe, all dialects basically preserve distinct forms for masculine, feminine and

Table 9. Variation in case marking, and plural forms of the ezafe

	Stand.K	SEK	SK	NWK
plural ezafe	-ên	-êd/-êt	-ê	-ê
sg. masc. oblique marking	vowel-raising /suffixation	suffixation	vowel-raising	vowel- raising / none

plural (though in NWK there is some collapsing of the system), and this remains one of the major features that distinguishes all varieties of Kurmanji from Sorani. However, there are differences in the form of the plural ezafe across different varieties of Kurmanji. Table 9 shows the distribution of the

¹⁶ A further unique feature of SEK is an additional ending in 3rd person singular *-itin*. We currently have no explanation for this form.

oblique marking of masculine singular nouns, and different forms of the plural *ezafe*:

Differences in the TAM system

Some dialects have grammaticalised particular tense, aspect and mood categories, creating important differences in the organisation of the grammar. For example, while standard Kurdish does not have a grammatical expression of progressive aspect (i.e. indicative present is used for all aspectual values), SEK has developed a specific construction for progressive, involving an *ezafe* after the clausal subject, e.g. *ež-ê nan-î di-xo-m* (I-EZ.MASC.SG food-OBL.MASC.SG IND-eat.PRS-1SG) ‘I am eating food’. Further features which show categorial and formal differences in tense-aspect-mood system of the dialects are different forms of the future particle (see Unger, this issue), the use of *ezafe* in combination with participles for forming present perfect tense, the presence of a past habitual particle *da* (Unger, this volume), and a clitic marker for indicating clauses expressing intention. The distribution of these features across the dialects is shown in Table 10.

Table 10. Differences in the tense-aspect-mood system of dialects

	Stand.K	SEK	SK	NWK
progressive aspect	-	+	-	+
present perfect with <i>ezafe</i>	-	+	-	-
future tense particle	dê/=ê	dê	=ê	=ê
past habitual particle	-	da	-	-
intention/prospective particle	-	-	-	=ke

Adpositions

Kurmanji dialects vary considerably in the kinds of adposition that are used in various functions. For example, the indirect objects of verbs such as ‘say’ are marked in SEK by a verbal particle *-e*, and they occur post-verbally. In NWK and SK, on the other hand (as representative of Central and Western Kurmanji respectively), indirect objects are marked by circumpositions such as *xi ... re/ra* ‘to’, and they occur pre-verbally. In the northern and westernmost dialects, the circumposition is often reduced to just the postpositional element. Note, however, that all dialects also permit indirect objects to be simply placed after the verb, with no adposition (this is in fact the commonest option with the verb ‘give’). A similar pattern is found with expressions of ‘with’: in SEK a preposition (*li*)*gel* is used, while in NWK and SK a circumposition as *bi ... re/ra* is used. The complete absence of these two circumpositions (*xi ... re* and *bi ... re*) in SEK constitutes a major difference setting off SEK from the rest of Kurmanji, and uniting it with Central Kurdish.

Proximal clitics

The demonstrative system of Stand.K includes determiners, as *ev* “this” and *ew* “that, those” together with their oblique versions *vi/vē* “this”, *wi/wē* “that”, *van* “these” and *wan* “those”. Most of the Kurmanji dialects, however, use also a clitic element attached to the proximate nouns (e.g. *ev kîṭēb* “that book” vs. *ev kîṭēb-e* “this book” and *ev kîṭēb-ene* “these books”), which has different singular and plural forms, as below:

	Stand.K	SEK	SK	NWK
Proximal clitic	-	-e/-ene	-	-a/-ana

Plural indefiniteness suffix

While the singular indefiniteness suffix *-ek* is shared across Kurmanji, the plural indefinite suffix *-in* is seen principally (if not exclusively) in SK, differentiating it from both SEK and NWK. This is one of the features that was adopted into Standard Kurdish, at least in codified grammars of the written form, but it is interesting to note that it has in fact a narrow distribution in the spoken language.

Conclusion

Among the regional varieties considered here, the most divergent is SEK, encompassing what is generally known as Badini. It has the lowest number of lexical items in common with the other dialects, and also shows the highest number of divergent features in the morphosyntax. In a number of respects, SEK reflects its proximity to Sorani, and some of the differences to the rest of Kurmanji could be explained through Sorani influence. However, that is not the whole story, because SEK also shows typical features of Kurmanji (see below), including very robust case marking of the oblique case, consistent maintenance of gender distinctions, and the aspirated/non-aspirated distinction on voiceless obstruents. These three features (among others) are completely lacking in Sorani, and thus speak against a view of Kurdish that would see Sorani and Kurmanji as simply two ends of an unbroken pan-Kurdish dialect continuum. The four varieties of Kurmanji considered here, however, can be considered to form a fairly typical dialect continuum, each fading into the areally contiguous region with no obvious sharp boundaries.

Looking a little closer at SEK, we can discern a number of features in the morphology that can reasonably be considered rather archaic: the richer set of verbal agreement markers (cf. Table 8), the more consistent use of affixal case-marking and gender, additional subjunctive forms lost in many of the other dialects, the consistent application of the ergative construction, and the use of non-canonical subjects for certain kinds of expressions (cf. Haig and Öpengin, forthcoming). In general, these are features that bespeak of a relatively stable speech community, with a long history of unbroken transmission and the comparatively small impact of adult second-language learners, who generally tend to simplify morphosyntax (see McWhorter, 2005; Trudgill,

2011 on contact effects and simplification in morphosyntax). The other dialects, on the other hand, are all characterised to one degree or another by loss of these richer morphological features. It is therefore conceivable that the other dialects, over the course of the northwestern expansion of Kurmanji speakers into Anatolia, became a lingua franca for speakers of other languages, in particular Neo-Aramaic, Arabic, and Armenian, and the impact of these “shifters” was an overall decrease in morphological complexity of the Kurmanji spoken in these regions. This remains of course speculative, but the presence of the additional aspiration contrast on stops and affricates in Kurmanji, particularly of Central Anatolia, is an absolutely typical result of a shift scenario, in this case from Armenian.

A simplified view of Kurmanji would recognise three main regional varieties: SEK, SK and NWK, the three that are the most distinct from each other. The other two are less clearly definable, and can probably be considered transitional zones: NK exhibits features of both NWK and SK, while SWK is also a transitional region between NWK and SK.

The geographically intermediary position of SK is also paralleled linguistically in that it shares features both with SEK and NWK. SK seems also to have gone through a substantial amount of morphosyntactic simplification (e.g. oblique marking, ezafe forms, Tense-Aspect-Mood system), which is hardly surprising if we consider that the speech zone of the dialect is geographically mostly flat (i.e. there are few natural barriers to inhibit mobility) and that over centuries, the dialect must have served as the lingua franca for non-Kurdish speaking speech communities in the region (Arabic, Aramaic and Armenian; see Noorlander, this issue, on Kurdish influence on Neo-Aramaic). In the same vein, it is also the dialect which exhibits most Semitic influence, especially in its lexicon but also in verbal morphology and phonology.

At the other end of the dialect continuum there is NWK, the Kurdish spoken in regions such as Elbistan. This dialect shows a number of highly divergent features, but it has been almost entirely neglected in previous research, and is sometimes even stigmatised among educated Kurds. We have only noted some of its features here (see Doğan, 2003 for some additional information), and can offer no convincing hypothesis for the high number of divergences in this dialect in comparison to the others, apart from the relatively isolated position of this dialect group within the totality of the Kurdish-speaking world. More research on NWK is an urgent desideratum within Kurdish linguistics.

References

- Anonby, E. (2003). Update on Luri: How Many Languages? *Journal of the Royal Asiatic Society*, 13(2), 171–97.
- Anonby, E. (2004/5). Kurdish or Luri? Laki's Disputed Identity in the Luristan Province of Iran. *Kurdische Studien*, 4/5, 7–22.

- Auer, Peter. (2005). The Construction of Linguistic Borders and the Linguistic Construction of Borders. In M. Fillpula, J. Klemola, M. Palander and E. Penttilä (eds.), *Dialects across borders* (3-30). Amsterdam: Benjamins.
- Bedir Khan, E. Dj. and Lescot, R. (1991[1970]). *Grammaire Kurde (Dialect kurmandji)*. Paris: Jean Maisonneuve.
- Blau, J. (1975). *Le Kurde de 'Amādiya et de Djabal Sindjār: Analyse Linguistique, Textes Folkloriques, Glossaires*. Paris: Klincksieck.
- Campbell, L. (1998). *Historical Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- Campbell, L. and Mixco, M. J. (2007). *A Glossary of Historical Linguistics*. Edinburgh: Edinburgh University Press.
- Dixon, R. M. W. (1997). *The Rise and Fall of Languages*. Cambridge: Cambridge University Press.
- Dorleijn, M. (1996). *The Decay of Ergativity in Kurdish*. Tilburg: Tilburg University Press.
- Doğan, M. (2003). Yılıya Bırakılmış Bir Dilin Çığlığı [The Cry of a Language Left to Destitution]. In M. Kömür (ed.), *Taw Dilo: Akçadağ-Elbistan Ağzıyla Kürt Folkloriünden Seçmeler* (14-26). Istanbul: Weşanên Enstîtuya Kurdî ya Stenbolê.
- Dyen, I., Kruskal, J. & Black, P. (1992). An Indoeuropean classification, a lexicostatistical experiment. *Transactions of the American Philosophical Society*, 82, 1-132.
- Enstîtuya Kurdî ya Parîsê. (2010). *Kurmancî: Rojnameya Taybetî ya Enstîtuya Kurdî ya Parîsê li ser Pirsên Zaravayê Kurmancî* [Kurmanji: Special Newspaper of Kurdish Institute of Paris on the questions of Kurmanji Dialect]. Istanbul: Avesta.
- Fox, A. (1995). *Linguistic Reconstruction: An Introduction to Theory and Method*. Oxford: Oxford University Press.
- Girdenis, A. and Mažiulis, V. (1994). Baltu kalbu divergencine chronologija. *Baltistica*, XXVII(2), 4-12.
- Gooskens, C. (2007). The contribution of linguistic factors to the intelligibility of closely related languages. *Journal of Multilingual and Multicultural Development*, 28/6, 445-467.
- Haig, G. (1998). On the interaction of morphological and syntactic ergativity: lessons from Kurdish. *Lingua*, 105, 149-173.
- Haig, G. and Öpengin, E. (forthcoming). Kurdish in Turkey: An Overview of Grammar, Dialectal variation and Status. In C. Bulut (ed.), *Minority Languages in Turkey*. Wiesbaden: Harrassowitz.
- Haspelmath, M. and Tadmor, U. (eds.) (2009). *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: Mouton De Gruyter.
- Hawar. (1998). *Hawar I, Hejmar 1-23 (1932-1933)*. Stockholm: Nûdem.
- Heggarty, P. (2012). Beyond Lexicostatistics: How to Get More out of "Word List" Comparisons. In S. Wichmann and A. P. Grant (eds.), *Quantitative Approaches to Linguistic Diversity: Commemorating the Centenary of the Birth of Morris Swadesh* (113-137). Amsterdam: John Benjamins.

- MacKenzie, D. N. (1961). *Kurdish Dialect Studies Vol. I*. London: Oxford University Press.
- MacKenzie, D. N. (1962). *Kurdish Dialect Studies Vol. II*. London: Oxford University Press.
- Mahmoudveysi, P., Bailey, D., Paul, L. & Haig, G. (2012). *The Gorani Language of Gawraju (Gawrajuyi), a Village of West Iran. Texts, Grammar and Lexicon*. Wiesbaden: Reichert.
- Kahn, M. (1976). *Borrowing and Regional Variation in a Phonological Description of Kurdish*. Ann Arbor, Michigan: Phonetics Laboratory of the University of Michigan.
- Kömür, M. (2003). *Taw Dilo: Aqçadağ-Elbistan ağzıyla Kürt folkloründen seçmeler [Sorrowful Heart: Selections from Kurdish Folklore in Aqçadağ-Elbistan variety]*. Istanbul: Weşanên Enstîtuya Kurdî ya Stenbolê.
- Labov, W. (1994). *Principles of Linguistic Change. Vol. 1: Internal factors*. Oxford: Blackwell.
- Le Coq, A. (1903). *Kurdische Texte. Kurmānġi-Erzählungen und -Lieder nebst einer Zāzā-Erzählung. Gesammelt und herausgegeben in der Urschrift und in Transkription*. Berlin [reprint 2003 Apa-Philo Press, Amsterdam]
- McWhorter, J. H. (2005). *Defining Creole*. Oxford: Oxford University Press.
- Preston, D. (2003). Language with an Attitude. In J.K. Chambers, P. Trudgill and Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change* (40–66). Oxford: Wiley-Blackwell.
- Prokić, J. (2010). Families and Resemblances. (Unpublished doctoral dissertation). University of Groningen.
- Prokić, J. and Cysouw, M. (2013). Combining Regular Sound Correspondences and Geographic Spread. *Language Dynamics and Change*, 3(2), 147–168.
- Ritter, H. (1971). Kurmānci-Texte Aus Dem Tûr'abdin. *Oriens*, 21–22, 1–135.
- Ritter, H. (1976). Kurmānci-Texte Aus Dem Tûr'abdin, Teil II. *Oriens*, 25–26, 1–37.
- Smith, J. and Durham, M. and Fortune, L. (2009). “Mam, ma troosers is fa'in doon!” Community, caregiver and child in the acquisition of variation in Scottish dialect. *Language variation and change*, 19(1), 63–99.
- Swadesh, M. (1955). Towards Greater Accuracy in Lexicostatistic Dating. *International Journal of American Linguistics*, 21(2), 121–37.
- Swadesh, M. (1959). Linguistics as an instrument of prehistory. *Southwestern Journal of Anthropology*, 15, 20–35.
- Trudgill, P. (2011). *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.
- Özsoy, A. S. and Türkyılmaz, Y. (2006). Front Rounded Vowels in the Sinemili Dialect of Kurmanji – A Case of Language Contact? In L. Johanson and C. Bulut (eds.), *Turkic-Iranian Contact Areas Historical and Linguistic Aspects* (300–309). Wiesbaden: Harrassowitz.
- Öpengin, E. (2012). Sociolinguistic Situation of Kurdish in Turkey: Sociopolitical Factors and Language Use Patterns. *International Journal of the Sociology of Language*, 217, 151–80.

- Jastrow, O. (1977). Zur Phonologie Des Kurdischen in Der Türkei. *Studien Zur Indologie Und Iranistik*, 3, 84–106.
- McCarus, E. N. (2009). Kurdish. In G. Windfuhr (ed.), *The Iranian languages* (587–633). London: Routledge.

Appendices

Appendix A: The basic vocabulary list

The first 100 items in this list are from the Leipzig-Jakarta list while the remaining 59 items are added within the frame of this research. The transcription is phonemic, and follows the conventions outlined above (on p. 146). The abbreviations (m) and (f) stand respectively for masculine and feminine grammatical gender found in all Kurmanji dialects (though this piece of information has not been noted for all of the items in the list), while (dir) and (obl) stand respectively for the direct and oblique case forms of the pronouns. The abbreviation (n.a.) indicates that the relevant dialect data is not available for the current analysis.

Item no	Meaning	SEK (Şemdinli)	SK (Midyat)	NK (Varto)	SWK (Adıyaman)	NWK (Elbistan)
1	fire	agir (m)	agir (m)	agir (m)	ar (m)	or (m)
2	nose	dîfin (f)	bêhvil (f)	poz (m)	poz (m)	poz (f)
3	to go	çûn	çûn (f)	çûyîn	çûn	çûyîn
4	water	aw (f)	av (f)	av (f)	av (f)	ov (f)
5	mouth	dew (m)	dev (m)	dev (m)	dev (m)	dov (m)
6	tongue	şezman (m)	ziman (m)	ziman (m)	ziman (m)	zimân (m)
7	blood	xîn (f)	xwîn (f)	xûn (f)	xûn (f)	xûn (f)
8	bone	hêstik (m)	hêstû (m)	hestî (m)	hestî (f)	hasî (m)
9	2SG	tu/te	tî/tê	tu/te	tu/te	tu/tê
10	root	rih (m)	ra (m)	k ^h ok (f)	k ^h ok (f)	k ^h ok (f)
11	to come	hatin	hatin	hatin	hatin	hôtin
12	breast	sîng (m)	sîng (f)	sîng (m)	sîng (m)	sîng (f)
13	rain	baran (f)	baran (f)	şîrî (f)	baran (f)	bôrân (f)
14	1SG	ez (dir)/mi(n) (obl)	ez (dir)/min (obl)	ez (dir)/mi(n) (obl)	ez (dir)/mi(n) (obl)	ez (dir)/mi(n) (obl)
15	name	naw (m)	nav (m)	nav (m)	nav (m)	nov (m)
16	louse	hispi (f)	speh (f)	sipi (f)	spi (f)?	ispî (m)
17	wing	p ^h er (m)	p ^h er (m)	p ^h er (m) bask (m) (human)	pîl (m)/bask (m)	pîl (f)
18	flesh/meat	goşt (m)	goşt (m)	goşt (m)	goşt (m)	goşt (m)
19	arm/hand	mil/dest (m)	çeng-mil (m)/dest (m)	p ^h i (m) mil (m)/dest (m)	p ^h i (m)/dest (m)	p ^h i (m)/dæst (m)
20	fly	mêş (f)	mêş (f)	mêş (f)	mêr (f)	mêş (f)
21	night	şew (f)	şev (f)	şev (f)	fav	şav (f)
22	ear	guh (m)	guh (m)	go (m)	guh (m)	guç/guçik (m)
23	neck	sukur (f)	sukur (f)	histû (m)	histî (m)	ustî (m)
24	far	dûr	dûr	dûr	dûr	dûr
25	to do/make	kirin/çêkirin	kirin/çêkirin	kirin/çêkirin	kirin/çêkirin	kirin/çêkirin
26	house	mal/xanî (m)	mal (f)/xanî (m)	mal (f)/xanî (f)	mal (f)/xanî (m)	xonî (m)/møl (f)
27	stone/rock	ber (m)	kevir (m)	kevir (m)	kævîr (m)	kævîr (m)
28	bitter	tûz	tûz	tûz	tûz	tuž
29	to say	gotin	gotin	gotin	gotin	gotin
30	tooth	didan (m)	dînan (m)	dîran (m)	dîran (m)	didon (m)
31	hair	p ^h irç (f)	p ^h ov (f)/p ^h or (m)	p ^h or (m)	p ^h or (m)	p ^h or (m)
32	big	mezin	mezin	mezin	mezin	gir
33	one	(y)êk	yek	yek	yêk	êk
34	who?	k ^h e	k ^h i/k ^h e	kî	k ^h i/k ^h e	k ^h i/k ^h e
35	3SG	ew/(e)wî-(e)wê	ew/wî-wê	ew/wî-wê	hew/wî - wê	æw/(æ)wî - (æ)wê
36	to hit/beat	lêdan	lêxistin	lêxistin	lêxistin	lêxistin
37	leg/foot	ling/pê (m)	ling-p ^h aq (f)/p ^h i-ling (m)	çip (m)/ling (m)	nig (m) (pied/leg)	nig (foot) (m)
38	horn	şax (f)	stîrh (m)	qoc (m)	qoc ^h (m)	ustru
39	this	ew ... e	ev/ew-that	ev-va/ew	va - vî/hew	æv/æv
40	fish	masî (f)	masî (m)	meşî	masî (m)	môsî (m)

REGIONAL VARIATION IN KURMANJI

41	yesterday	duhu(nê)	duho	do	dihu	do
42	to drink	xarinewe	vexwarin	(ve)xwarin	vexwarin	vexorin
43	black	reš	reš	reš	reš	reš
44	navel	nawik (f)	navik (f)	navik (f)	nêvk (f)	nâvik (f)
45	to stand	rawestan	sekinandin	sikinandin	Sekinîn	sêkinîn
46	to bite	leq dān	gez kirin	kit kirin/ gez kirin	gez kirin	gez kirin
47	back	p ^h ist (f)	p ^h ist (f)	p ^h ist	p ^h ist	pōš
48	wind	ba (m)	ba (m)	bā (m)	ba (m)	bō (m)
49	smoke	dūk ^h el (f)	dūman (f)	dū (m)	dū (m)	tu/tuman (f)
50	what?	ç ^h i	ç ^h i	ç ^h i	ç ^h i	ç ^h i
51	child (kin term)	biçûk	zafoke (f)	zarû/zar	zar/zark (m)	dêl (m)
52	egg	hêlik/hêk (f)	hêk (f)	hêk	hêk (f)	hêk o -f
53	to give	dan	dan	dayîn	dan	doyîn/don
54	new	nû	nû	teze	nô	no
55	to burn (intr.)	sotin	şewitandin	şewitandin	şewitandin	şewitandîn
56	not	ne	na	na	na	nō
57	good	baş	baş	rind/baş	rind	rind
58	to know	zanîn	zanîn	zanîn	zanîn	zōnîn
59	knee	ç ^h ok (m)	k ^h abok (m)	ç ^h ok (m)	ç ^h ong (f)	ç ^h ok (m)
60	sand	xîz (m)	qûm (f)	qûm (f)	qûm (f)	qûm
61	to laugh	k ^h enîn	k ^h en (m)	k ^h enîn	k ^h enîn	k ^h enîn
62	to hear	bihistin/hay jê bûn	his kirin/ hisandin	bihistin	bihistin	bistin
63	soil	ax (f)	ax (f)	xwelî (f)	ax (f)	ox (f)
64	leaf	pelik/belg (m)	p ^h el (f)	p ^h el (f)	p ^h elç (f)	çilo (f)
65	red	sor	sor	sor	sor	sur
66	liver	cerg (m)	kezeb (f)	ciger (f)	kezew (f)	çigar/kezew (m)
67	to hide	şartinewe	veşartin	veşartin	veşartin	veşartin
68	skin/hide	çerm (m)	çerm (m)	çerm (m)	çerm (m)	çerm/post (m)
69	to suck	mêtin	mižandin	mižandin	mižîn/ mižandin	mitîn/ mižîn
70	to carry	k ^h êşan	hîlgirtin/ kişandin	k ^h işandin	k ^h işandin	k ^h işandin
71	ant	mêrû (f)	mîro (f)	mirjolek (f)	mori (m)	gêra (m)
72	heavy	giran	giran	giran	giran	giran
73	to take	birin	birin	birin	birin	birin
74	old	kewn	kevin	kem	kêvin	kevin
75	to eat	xarin	xwarin	xwarin	x ^w arin	xorin
76	thigh	qalç ^h e (f)	k ^h elef (f)	hêt (f)	qalç ^h e (f)	qalçe (f)
77	thick	ştûr	qalîn/ştûr	qalînd	qalîng	qalîng
78	long	dirêz	dirêz	dirêz	dirêz	dirêz
79	to blow	p ^h if kirin	p ^h eqandin	p ^h if kirin	p ^h ifê kirin	p ^h uf kirin
80	wood	dar (f)	dar (f)	dar (f)	dar (f)	dor (m)
81	to run	bezîn	baz dan	revîn	revîn	boz dan
82	to fall	ketin	ketin	ketin	ketin	ketin
83	eye	ç ^h aw (m)	çaşv (m)	çeşv (m)	çav (m)	çov (m)
84	ash	xûlî (f)	xwelî (f)	xwelî (f)	ari (f)	x ^w elî
85	tail	dûlik (f)	dûv (m)	boçik (f)/dêl (f)	boç (f) dû	doçik (m)
86	dog	şe (m)	şe	kûçik (m)/se (m)	kûçik (m)	so/kudik
87	to cry/weep	giriyan	giri	giri	giri kirin	giri
88	to tie	girêdan	girêdan	girêdan	girêdan	girêdan
89	to see	dîtin	dîtin	dîtin/dîn	dîn	dîn
90	sweet	şîrîn	şîrîn	şîrîn	Şîrîn	şîrîn
91	rope	bendik (m)/k ^h indir (f)	ta (m)/wêris (m)	bend (m)	ben (m)/ kêndir (f)	bên (m)/ kindir (f)
92	shade/shadow	sîber (f)	sîh (f)	sî (f)	sî (f)	sî (f)
93	bird	cûçik (f)	çûk (m)	çûçik (f)	çûk (f)	çêçûk (f)
94	salt	xê (f)	xwê (f)	xwê (f)	xwê (f)	xwê (f)
95	small	kiç ^h ke	piçûk	piçûk	çûçik/çûk	çûk
96	wide	fire	fêreh	fire	fere	fere
97	star	stêrk (f)	stêrik (f)	histêrik (f)	hêstirk (f)	istewrik (m)

98	in	di .. da/(di) naw .. da	di .. da/li hindurū	di (hindirē) .. da	hundir/di ..da	hundirī/dæ .. dō
99	hard	req	hişk	ʕeşk	sert	seft
100	to crush/grind	p ^h elēqandin	p ^h irē ^h iqandin	p ^h elē ^h iqandin	p ^h elē ^h iqandin	pēnandin
101	daughter/girl	kič ^h (ik) (f)	kečik/keč (f)	qiz/keč	kečik (f)	kəčik (f)
102	lie	dirū (f)	deŕū (f)	direw (f)	derew (f)	derew (f)
103	sky	ʕesman (m)	azman (m)	ʕezman	ʕezman	əzmən (m)
104	sun/day	rož (f)	ro/rož (f)	ro(ž) (f)	ro/ro (f)	ro/rū (f)
105	morning	spēde (f)	sibeh (f)/ ševēq (f)	sibē (f)	siwē (f)	siwē/sibē (f)
106	moon	hewī (f)	heyv (f)	hiv (f)	hēv (f)	hiv (f)
107	man	miro(v) (m)/f	zīlam, mēr, īnsan	mērik/meriv	mēr/merī	mēr/mæri
108	arrow	kūwān (f)	kevan (f)	kevan (f)	kevan (f)	kuvan
109	all	hemi	hemi	gi/gišt/hemū	gičik	gišt
110	here	ēre	li vir/ li vē dē/ev der	li vira/vira	li vir/virē	læ vir/æv dærē
111	there	were	li wir/wē dē/ew der	li wir/wē dera/wira	li wir(ē)/wira	læ wē/læ wē derē/æw/wē derē
112	2PL	hing	win	hūn	hūn	hūn
113	now	nuke	niha	niha/nika	niha	wēstō
114	woman	žin/žinik (f)	pīrek (f); jin (f)	žin	žin	žin(ik)
115	today	ewro	iro	iro	hiro	hiro
116	if	ku/heke	go/eger	a qas/a - waqa	ko/eger	ko/əgər
117	sister	xüšk (f)	xweh (f)	xweng	x ^w aying	xüčik/x ^w ong
118	wolf	gurg (m)	gur (m)	gur (m)	gur (m)	gur (m)
119	beard	fidin (f)	rih (f)	rū (m)	rihi	rū (m)
120	easy	senahī	hesanī	rihet	qole	rihat
121	sheep	mihī (f)	mih (f)	pez (m)	mihi (f)	mī (f)
122	one	yēk	yek	yek	yēk	ek
123	two	du	du/dido	du/dudu	didu/du	(du)du
124	three	sē	sē/sisē	sē/sisē	sisē/sē	(si)sē
125	seven	hewt	heft	heft	heft	haft
126	eight	hešt	hešt	heyšt	heyšt	hašt
127	nine	nehe	neh	neh	nehe	na
128	ten	dehe	deh	deh	dehe	da
129	eleven	yanzde	yazdeh	yonzde	de w yek	da w yek
130	twelve	dānzde	duwanzdeh	donzde	de w didu	daw dudu
131	seventeen	hevde	hivde	hivde	de w heft	daw u haft
132	thirty	sih	sī	sī	sī	sī
133	fourty	č ^h il	č ^h el	č ^h il	č ^h el	č ^h al
133	fifty	p ^h ence	pēncī	pēncī	pēncī	pēncī
134	ninety	nehwērt	nod	nod	nod	nod
135	hundred	šed	šed	sid	sed	sed
136	thousand	hizar	hazar	hezar	hezar	hazær
137	turtle	k ^h üsel (m)	k ^h iso (m)	k ^h usi (f)	reç - reqesēl (m)/f	k ^h üsē (m)
138	work	šol (f)	šoXil (m)	iš (m)/šixol (m)	šuxil (m)	šoXul (m)
139	end	dūmayik	dawī	dawī	sonXī (f) pašī (f)	song
140	hedghog	žūži (m)	žižo (m)	šizū (f)	žūži (m)	žuži (m)
141	eyebrows	mižilank (f)	mižgūl	bižang	bižang (f)	biži (f)
142	leg	ling (m)	ling	č ^h ip (m)	nig (m)	žuni (f)
143	finger	tipl (f)	tilih (f)	t ^h il	pēči (f)	pēči/t ^h ili
144	rooster	diqil (m)	diik (m)	diik (m)	diik (m)	diik (m)
145	owner	xudan (f)/m	xwedī	xwedī (m)	x ^w adi	x ^w odi
146	father	bab (m)	bav	bav	bav (m)	bōv (m)
147	thirsty	t ^h ēni	tī	tī	tī	tī
148	hungry	birsī	beŕči	birči	birči	birči
149	how	č ^h itū	čawa	č ^h ito	č ^h ita	č ^h ito
150	how	k ^h utū	čeŕ	č ^h er	(n.a.)	č ^h ir
151	that way	wet ^h ū	halo	aha	ha/hani/hana	wer
152	thus	hū ^h tū	wūlo	wer	werga/werganī	ho/hoyna
153	that	hind	evqas/wilqas	a qas/a - waqa	wiqes/wiqes	vaqesi

REGIONAL VARIATION IN KURMANJĠ

	much/that many					
154	then	hingî	hingî	wî çaxî/ hingê	wê ç ^h ağê	wêngê(î)/wê ç ^h ağê
155	as ... as	hindî	hingî	qasê ..	hîn	qasa ..
156	like	wekû/wekî	wekê	mîna/ fena	mîna/nola	mînê/wekî/no lo
157	next to, by	lali/li nik	(li) cem/(li) ba	li hinda/li kêleka min/ba/ cem	li cem/li ba/li hinda	řev/hindê
158	side	qerāğ (f)	kevî/k ^h êlek/ ře x	kêlek	kêlek	k ^h enêr
159	yet	hêřta	hê	hîna	hîn	hên

Appendix B: Sample list of vocabulary items showing high phonological variation

Item no	Meaning	SEK	SK	NK	SWK	NWK
1	father-in-law	xezür	xezür	xezür	xezür	xenzür
2	walnut	güz	güz	güz	güz	guz
3	far	dür	dür	dür	dür	dür
4	eye	ç ^h aw	çaşv	çeşv	çav	çov
5	snake	mar	mar	meşr	meşr	mər
6	bitter	t ^h al	tehl	teşl	teşl	təl
7	flat	p ^h an	pehn	peşn	peşn	pən
8	sister	xüşk	xweh	xweng	x ^w eyîng	xüşk
9	salt	xê	xwê	xwê	x ^w ê	x ^w ê
10	to read	xandin	xwendin	xwendin	xandin	x ^w andin
11	small	biçûk	piçûk	piçûk	çûçik/çûk	çûk
12	a little	piçek	hindik/piçek	hindik/piçek	hindik	çeki/hindik
13	it's possible	dibît	dibê	diwe	diwe	debi
14	s/he says	dibêjît	dibêjê	diwê	diwê	debê
15	because of	(ji)be(r)	ji boy wilo	se-ba/sewa	sêwa wê/me?na wê	sawê/sawa (te)
16	kabab	kebab	kebab		kiwaw (f)	kebab
17	one unit/single	heb/hebekê	heb	heb/hew	hewek	hawek
18	water	aw	av	av	av	ov
19	night	şew	şev	şev	şav	şav
20	to come	hatin	hatin	hatin	hatin	hötin
21	to know	zanîn	zanîn	zanîn	zanîn	zönin

Appendix C: Sample list of vocabulary items showing high lexical variation

Item no	Meaning	SEK	SK	NK	SWK	NWK
1	lost	bezir	wunda	wenda	wenda	wəndə
2	uncle's wife	žinmam (f)	jinam	jinap	amožin	əmož/xəlož
3	father's brother	mam (m)	am	ap	ap	əp
5	mirror	qūdi (f)	ŕeynik	ŕeyne	neyk (f)	neynik (f)
6	wet	teŕ	šil/tir (dar)	šil/ter	šil/ter	šil/tēn
7	stable	hel/p ^h age (f)	kox (f)	tewle (f)	axir (f)	əxir (m)
8	cradle	landik (f)	dergūš (f)	bēšik (dergūš=be bek)	dergūš (f)	dərgūš (f)
9	throat	hewk (f)	qirik (f)	qirik (f)	qirik (f)	gəwri (f)
10	watermelon	šimti (m)	šebeš (m)	zeveš	zeweš (m)	qerphūz
11	melon	gundore (f)	p ^h etix (m)	qawin	k ^h elek (m)	xx
12	potato	sēwik (m)	p ^h etat (f)	kartol (f)	p ^h ateti z (f)	p ^h etik (f)
13	tomato	t ^h emate (f)	bajan (f)	t ^h ematēs	fireng (f)	t ^h emutos/t ^h em irtos (f)
14	to understand	tēgehištin	fēm kirin	feŕm kirin	feŕm kirin	fōm kirin
15	to learn	fēr būn	ŕelimandin	ŕelimandin	hū(n) būn/h ūlikē būn	belli kirin
16	to be able to	šiyān	karin	karin	kanin	šə kirin
17	to finish	biriyan/xilas būn	xelas būn/qedehan/	xilas būn	xelas būn	t ^h ewa būn
19	to collect	xir kirin/kom kirin	dan hev dū/dan hevdū/	to v ki-rin/kom kirin	t ^h op ki-rin/t ^h o pa hev kirin	beref kirin
20	to spend	xerj kirin	xerj kirin	xerj kirin	xerj kirin	xerj kirin
21	to send	hinartin/ŕē kirin	šiyandin	šandin	šandin	šandin
22	to look at smt/smb	berē xo dan/lē westan	mēyzandin	mēze ki-rin/nēŕin	nēŕin	mēz kirin
23	to burn	sotin	šewitandin	šewitandin	šewita ndin	šəwitandin
24	to deceive	lēbandin	xapandin	xapandin	xapandin	xəpandin
25	to sleep	niwistin	ŕazandin	ŕaketin	ŕaketin	rəkətin/rəmə dīn
26	to speak	axiftin	šiteŕlandin	qise dan	deyn kirin	deyn kirin
27	to hold	helgirtin	ŕahištin/hilgirtin	hilgirtin	helgirtin	hilgirtin
28	to get tired	šeqi būn	betlandin	westandin	westi-yan	wəstiyən/osən būn
29	to wait for	xo čawri (yēkē) girtin	li hēviyē/benda būn	li hēviya/li benda	li bendē	pən

					man	
30	to walk	řē we čün	mešandin	řē va čün	meši- yan	mešin/yērmış bün
31	to want	wiyan/xwastin	viyan/restricted usage	xwastin	xwas- tin (no viyan)	xwəstin (no viyan)
32	to hang	helawēstin	daleqandin	darda kirin	bi dar xistin	dardəkirin
33	to help	hař kirin	ař kirin	alř kirin	n.a.	orř kirin
34	change	gořin	(n.a.)	degiřandin/ guheřandin	gihē- tin	g(uh)orřin (bigorřim)
35	squeeze	guwiřtin	(n.a.)	givařstin	guviřti n	giviřtin
36	reach	gehiřtin	(n.a.)	giřtin	(n.a.)	gēřtin

Appendix D: Verb paradigms

SEK verb paradigms					
Stand.K	Meaning	3SG Present Indicative	2SG Imperative	3SG Past	3SG Present Subjunctive
bün	'be'	(y)e	be	bü	bīt
čün/čüyin	'go'	dičīt	here	čū	bičīt
hatin	'come'	dihēt	were	hat	bihēt
kirin	'do'	diket	bike	kir	biket
birin	'take'	dibet	bibe	bir	bibet
dan/dayin	'give'	didet	bide	da	bidet
xistin	'drop'	diēxīt	bēxe	ēxist	biēxīt / bēxīt
ketin	'fall'	dikewīt	bikewe	ket	bikewīt
anin	'bring'	dīmīt	bīne	īna	bīnīt
xwarin	'eat'	dixot	bixo	xar	bixot
řazan	'sleep'	diřazēt	řāzē	řāzā	řāzēt
danin	'put'	dadinēt	danē	dana	danēt
gotin	'say'	dibēžīt	bēže	got	bežīt
đitin	'see'	dibīnīt	bibīne	đit	bibīnīt

175

SK verb paradigms					
Stand.K	Meaning	3SG Present Indicative	2SG Imperative	3SG Past	3SG Present Subjunctive
bün	'be'	(y)æ	bæ	bū	bē
čün/čüyin	'go'	dičē	here	čū	herē
hatin	'come'	tē	were	hat	werē / newē
kirin	'do'	dikē	bikæ	kir	bikē / nekē
birin	'take'	dibē	bibæ / nebæ	bir	bibē / nebē
dan/dayin	'give'	didē	bidæ	da	bidē
xistin	'drop'	dixē	bixæ	xist	bixē
ketin	'fall'	dikevē	bikevæ	ket	bikevē
anin	'bring'	tēnē / naēnē	bēne	anī	bēnē / neēnē
xwarin	'eat'	dixwē	bixwæ	xwar	bixwē
řazan	'sleep'	diřazihē	řaze	řaziha	řazē
danin	'put'	dideyne	deyne	deyna	deynē
gotin	'say'	dibēžē	bēže	got	bēžē
đitin	'see'	dibīnē	bibīne	đit	bibīnē
man	'stay'	dīmēnē	bīmēne	ma	bīmēne

NK verb paradigms					
Stand.K	Meaning	3SG Present Indicative	2SG Imperative	3SG Past	3SG Present Subjunctive
bûn	'be'	y/e	be	bû	be
çûn/çûyîn	'go'	t ^h eře	heře	çû	heře
hatîn	'come'	tê	were	hat	were
kirîn	'do'	dike	bike	kir	bike
birin	'take'	dibehere	beheře	bir çû	beheře/bibe
dan/dayîn	'give'	dide	bide	da	bide
xistin	'drop'	dixîne	bixîne	xist	bixîne
ketin	'fall'	dikeve	bikeve	ket	bikeve
anîn	'bring'	tîne	bîne	anî	bîne
xwarin	'eat'	dixwe	bixwe	xwar	bixwe
razan	'sleep'	radikeve	fakeve	raket	fakeve
danîn	'put'	datîne	dayne	danî	dayne
gotin	'say'	dibîne/diwîne	bibîne/biwîne	dît/dî	bibîne/biwîne
dîtîn	'see'	dibê/duwê	bibê/bêje	got/go	bibê/bêje

SWK verb paradigms					
Stand.K	Meaning	3SG Present Indicative	2SG Imperative	3SG Past	3SG Present Subjunctive
bûn	'be'	e/ye	be	wû	be
çûn/çûyîn	'go'	dare	here	çû	here
hatîn	'come'	tê	were	hat	were
kirîn	'do'	dike	bike	kir	bike
birin	'take'	diwe	biwe	bir	biwe
dan/dayîn	'give'	dide	bide	da	bide
xistin	'drop'	dixe	bixe	xist	xe
ketin	'fall'	dikeve	bikeve	ket	bikeve
anîn	'bring'	tîne	ûne	anî	ûne
xwarin	'eat'	dix ^w e	bix ^w e	x ^w ar	bix ^w e
razan	'sleep'	radikeve	fakeve	raket	fakeve
danîn	'put'	dîtîne	dîne	danî	dîne
gotin	'say'	diwê	biwê	got	biwê
dîtîn	'see'	diwîne	biwîne	dî	biwîne

NWK verb paradigms					
Stand.K	Meaning	3SG Present Indicative	2SG Imperative	3SG Past	3SG Present Subjunctive
bûn	'be'	æ / yæ	bæ	bû	bi
çûn/çûyîn	'go'	tæri	here	çû	heri
hatîn	'come'	tê	were	hət	weri
kirîn	'do'	dækæ	bikæ	kir	biki
birin	'take'	dæbæ	bæ	bir	bæ
dan/dayîn	'give'	dædi	bidæ	də	bidi
xistin	'drop'	têxi	bixæ	xist	bixi
ketin	'fall'	dækævi	bikævæ	kæt	bikævi
anîn	'bring'	tîni	wînæ	ənî	wîni
xwarin	'eat'	dæxæ / dæxwæ	bixwæ	xwær/xwər	bixwi
razan	'sleep'	radækævi	fakevæ	fakeæt	fakevi
danîn	'put'	dîtîni	dînæ	dənî	dîni
gotin	'say'	(ewî) debê	biwê	go(t)	biwê
dîtîn	'see'	dæbîni	biwînæ	dî	biwîni