Anju Saxena, Anna Sjöberg, Padam Sagar, and Lars Borin

4 Linguistic variation: a challenge for describing the phonology of Kanashi

Abstract: Kanashi exhibits a great deal of variation on several linguistic levels, which raises questions of a theoretical and methodological nature relevant to the formulation of useful and faithful linguistic descriptions of Kanashi. In this chapter, we address such questions in connection with working out a description of the phonology of Kanashi as part of a larger language documentation effort. Specifically, we discuss two aspects of the sound system of Kanashi where we have noted considerable variation among our language consultants, and which consequently necessitate reflection and discussion over their place – if any – in the phonology of Kanashi: aspirated voiced stops and geminated (long) consonants.

Keywords: Kanashi, Sino-Tibetan, phonology, phonetics, variation

Chapter overview:

1 Introduction — 131	4.1 Phoneme inventory: aspirated voiced stops — 136
2 Data collection — 132	4.2 Phoneme inventory: geminate consonants — 138
3 Linguistic variation in Kanashi — 133	5 Discussion — 139
4 Two instances of variation in the sound system of Kanashi — 135	6 Conclusions — 143

1 Introduction

In this chapter, we will describe some challenges we have faced in preparing a description of the phonology of Kanashi as part of a more extensive language-documentation effort among whose main results are the descriptions presented in Chapters 2 and 3 in this volume. The challenges are primarily caused by the

† The following notational conventions are used in this chapter. Phonetic transcriptions are given in IPA notation in square brackets “[...]”. Even though our transcription conventions for Kanashi are intended as phonemic, whenever we wish to stress that phonemes and phonemic representations are under discussion, we write single phonemes and phoneme sequences surrounded by “/.../”. We represent geminate consonants as doubled (biphonemic) rather than long (i.e. we write *daddu* rather than *dadu*), but we would like to stress that this does not imply a strong preference on our part for the one or the other analysis.

amount of inter- and intra-speaker variation evidenced in our recordings. We will discuss the following two phonological phenomena to illustrate our case, both concerning the phoneme inventory of Kanashi:

1. the status of aspirated voiced stop consonants (Section 4.1)
2. geminate ~ singleton consonant variation (Section 4.2)

2 Data collection

The following description is based mainly on the speech of three Kanashi speakers (2,063 sound files for 975 words): one older male Kanashi speaker (older male, “OM” in the following) and two younger speakers, one man (younger male, “YM”) and one woman (younger female, “YF”). At the time of the data collection OM was about 50 years old, YM around 22 years old and YF around 28 years old. The female speaker had received some formal education. Both male speakers were illiterate; they had not received any formal education. All three speakers were born and brought up in the village. Like most other Kanashi speakers, they leave the village occasionally. All three could understand Hindi (hin), the national language of India, and Kullu Pahari (kfx),¹ the locally dominant language, both Indo-Aryan (IA) languages. While the female speaker could speak Hindi quite well, the male speakers (especially YM) spoke a mixed Hindi with strong influence of their mother tongue Kanashi.

The Kullu Pahari data in this chapter is from our fieldnotes;² Hindi data is from McGregor (1993).

Apart from lexical items, we have also collected elicited phrases and sentences and some narratives. The elicitation was done in Hindi. All recordings were done with a general documentary purpose, that is, not with a particular phonetic experiment in mind. This means that in some cases the materials are not optimal for the analyses conducted on them (e.g. having list intonation, not having tokens of the same word from all speakers etc.). The acoustic phonetic analysis of the sound files was done using Praat (Boersma & Weenink 2018).

¹ It is also referred to as Kullui/Kulluvi/Inner Siraji.

² Thanks to our language consultants Mrs Kanta Devi and Mrs Meena Bodh for their input.

3 Linguistic variation in Kanashi

Even with the limited data available on Kanashi, the language exhibits a great deal of variation, at several linguistic levels. For instance, there are no less than four different ways of forming composite numerals in Kanashi for each cardinal in the range 21–99, and in addition the corresponding Hindi numerals can also be freely used instead. The Kanashi numerals are described and discussed in more detail in Chapter 5 of this volume.

IA loanwords are found in the language in both an older and a more recent form, showing different degrees of integration into the linguistic system of Kanashi. Lexical items belonging to the older loanword stratum have undergone various adaptation processes, not only in their pronunciation, but notably also addition of the adaptive suffixes *-es/-aŋ/-iŋ* to IA noun and adjective stems (see Table 1 and Chapter 6).

Table 1: Older IA loanwords in Kanashi

Gloss	Kanashi	Kullu Pahari
‘traditional guesthouse’	<i>marhaŋ</i>	<i>marh</i>
‘treasury’	<i>baŋda:raŋ</i>	<i>b^hanḍa:r</i>
‘cattleshed’	<i>k^huɽaŋ</i>	<i>k^huɽ</i>
‘name’	<i>na:maŋ</i>	<i>na:m</i>
‘wall’	<i>bitiŋ</i>	<i>bit</i>
‘bucket’	<i>ba:lṭiŋ</i>	<i>ba:lṭi</i>
‘thief’	<i>ʈforas</i>	<i>ʈfor</i>

Distinct from this, more recently borrowed IA nouns and adjectives occur without these adaptive markers, and also retain the phonological structure of IA to a greater extent than in the older loanword stratum, e.g. *dahi* ‘yoghurt’; *laṛi* ‘wife’.

This now leads to situations where we find lexical doublets going back to the same IA item in our data, e.g. *gra:maŋ* ~ *gra:m* ‘village’; *na:maŋ* ~ *na:m* ‘name’; *diliŋ* ~ *dilli* ‘Delhi’. This, too, contributes to variation – both in phonological realization and in creating lexical and morphological variants. In such cases it is not easy to draw a line between borrowings and instances of code mixing. For the most part, we have treated these cases as borrowing, with one main exception. In instances where we have pronunciation variants of a certain lexical item with IA etymology, e.g. [puḍa] ~ [puṭʃas] ‘fifty’ (cf. standard Hindi *pacās* ‘fifty’), we examined if the more standard IA form is found only in the speech of the literate speaker(s), and in those instances where this turned out to be true we chose the

other variant as our phonemic representation of that lexical item, e.g. *puḍa* ‘fifty’. Such variation is not unique to Kanashi, but instances of variation which cannot be correlated with some linguistic/socio-cultural factors, pose a challenge to the linguist, especially when the phoneme inventory is often seen as a closed set.

In this chapter we will describe some challenges we have faced in describing the phonology of Kanashi in the face of the variation found. It is likely that several factors contribute to the variation that we find in Kanashi. Kanashi is an oral language. It is in the early days of documentation, so that uncertainties about the phonemic status of segments may possibly translate into variation in transcriptions. Some of the variations that we see in our data may be due to variation among speakers correlating with demographic and sociocultural factors, e.g. literacy and increasing intrusion of Hindi (one of the two national languages of India and the official state language of Himachal Pradesh, as well as the medium of instruction in schools).

Another suggested factor contributing to variation in descriptions of lesser-known languages is (inadvertent) inclusion of more than one geographical variety into the description. This does not seem to be a relevant factor in the case of Kanashi, however, as Kanashi is spoken only in one village (Malana), and the physical structure of the village is rather compact with houses close to one another.

Further, there is the observation made about many endangered languages that there may be variation without any apparent explanation – no stylistic differences, no social variables such as age or class correlating with the variation (Cook 1989; Palosaari & Campbell 2011). This may be the result of previously obligatory phonological processes being lost, giving rise to free variation. The mechanisms behind these changes in phonology in endangered languages are probably multifarious. It has been suggested that a pivotal role is played by so-called semi-speakers, i.e. speakers who have not learnt the language fully (Cook 1989). In the case of Kanashi, children still acquire the language, there are no semi-speakers and the language is used in everyday life in the village. Thus, the role of imperfect learning or semi-speakers in any variation found in Kanashi should be negligible.

Others stress the influence of language contact, e.g. Andersen (1982), who proposes the generalization that contrasts are maintained that exist both in the target language and the language with which it is in contact and also contrasts that carry a high functional load. As Malana village where Kanashi is spoken has neighboring villages where IA languages are spoken, one assumption could be that linguistic variation in Kanashi can possibly be due to contact with speakers of neighboring villages. This, however, is not a major contributing factor, as Kanashi speakers, in their everyday life, maintain a clear distinction between Kanashi and non-Kanashi people.

The effects of language contact, however, cannot be discounted, as most Kanashi speakers are bilingual. Despite the focus on maintaining distance from non-locals, there is regular interaction for religious and economical purposes between Malana and the neighboring IA-speaking villages. As mentioned above, the Kanashi lexicon in fact contains a large share of IA loanwords. In recent times, with the advent of (satellite) television, mobile phones and the internet, the influence of Hindi has become an integral part of the linguistic environment of Kanashi speakers.

Finally, we should perhaps also not overlook the circumstance that we have from an early age been prescriptively exposed to our own standard language with its concomitant normative conceptual framework connected to language standardization, which easily could be unreflectingly carried over into language description conducted by us in our capacity of academic linguists, arguably a special case of “written language bias in linguistics” (Linell 2005). It is well recognized that a central aspect of the creation of standard languages is reduction of variation; some variants are simply excluded from the standard (Joseph 1987: 126ff). This means that a language like Kanashi would be expected to exhibit more variation simply by virtue of not being standardized.³

The fact remains that we find a great deal of variation in Kanashi, and more research is definitely needed in order to ascertain how (if at all) this variation correlates with demographic and other factors. Whatever its causes, the variation has consequences for the linguistic description of Kanashi, and as a concrete illustration of this we now turn to a discussion of two instances of variation in the sound system of Kanashi which have forced us to think about what should be counted as evidence for phonemic status of a phonetic segment or feature.

4 Two instances of variation in the sound system of Kanashi

Kanashi exhibits several kinds of variation in its sound system. For example, degree of final devoicing of voiced stops, aspirated voiced stops alternating with their unaspirated counterparts, varying vowel and consonant length, and some

³ The kinds of observations that we make here about Kanashi are far from new: “Variation in Navajo pronunciation had long disturbed Haile (to Sapir, 30 March 1931: SWL): ‘Sometimes I do wish that the informants would be more careful in pronunciation and follow some system which would conform to theory. ... Apparently no excuse, excepting that informants are too lazy to use it correctly.’” (Darnell 1990: 257)

aspects of vowel quality (in particular the presence of a vowel in the space between /a/ and /o/ in OM and a more centralized /u/, often approaching [ə], in YM).

Below, we will focus on two of these instances of variation, namely voiced aspirates and geminate consonants.

4.1 Phoneme inventory: aspirated voiced stops

For each place of articulation, many IA languages exhibit four series of oral stop phonemes with respect to their manner of articulation, characterized by presence or absence of voicing and aspiration. Thus, oral stop phonemes can be voiceless and unaspirated, voiced and unaspirated, voiceless and aspirated, and voiced and aspirated, e.g. /p/, /b/, /p^h/, and /b^h/.

Kanashi has stop phonemes at four places of articulation: bilabial, alveolar, retroflex and velar. For each of these places of articulation, there is an uncontested basic three-way distinction between unaspirated voiced stop, unaspirated voiceless stop and aspirated voiceless stop, as demonstrated by the near minimal triplet /du/ [3SG.DIST], /to/ ‘be’ and /t^ho/ ‘up’. The aspirated voiceless stops are somewhat limited in their distribution and are almost only found in word-initial position, with a few exceptions, but since they contrast with the other two stop series in that position, their phonemic status seems incontrovertible.

In our data we also find aspirated voiced stops – [b^h], [d^h], [d^h], and [g^h]. However, their distribution is such that it is unclear if they are to be considered parts of the Kanashi phoneme inventory. Voiced aspirates are not a characteristic feature of ST languages, although Matisoff (2003) notes that many Himalayish ST languages that are in close contact with IA languages borrow them, first in loanwords and then extending their use to the entire lexicon.

In Kanashi voiced aspirated stops are predominantly found in the IA part of the lexicon, see Table 2. Hindi is not necessarily the source language, but is here taken to represent IA. As we can see here, in all these examples we find both variants in our material – with and without aspiration.⁴ They appear to be in free variation. That is, lexical items which are realized with aspirated voiced stops are also realized with plain voiced stops, by the same speakers.

⁴ The second point is that voiced aspirates are, similarly to voiceless aspirates, limited in their distribution. They are almost only found in word-initial position (examples of exceptions to this in our data: [gind^hi] ‘ball’ and [prad^ha:n] ‘chieftain’). This, however, is a minor point, especially since several consonant phonemes in Kanashi that are clearly supported also show some distributional limitation.

Table 2: Aspiration alternation in voiced stops

Kanashi	Hindi	Gloss
[b ^h ukamp] ~ [bukamp]	<i>bhūkamp</i>	‘earthquake’
[b ^h au] ~ [bau]	<i>bhāi</i>	‘brother’
[g ^h ari] ~ [g ^h aɾi] ~ [gaɾi]	<i>ghaṛī</i>	‘clock, watch’
[g ^h oɾa] ~ [goɾa]	<i>ghoṛā</i>	‘horse’
[d ^h auga] ~ [dauga]	<i>dhāgā</i>	‘thread’

If we would like to maintain that a morpheme should normally have *one* unitary underlying shape (a sequences of phonemes), this raises the question of which phoneme should be taken as being present in the underlying form of these lexical items: the aspirated voiced stop, contrasting with the voiced unaspirated stop, or the voiced unaspirated stop, or possibly some underspecified segment or (equivalently) an archiphoneme? These three alternatives are represented below, together with the resulting underlying representation for the word for ‘brother’.

Alternative 1	Alternative 2	Alternative 3
/b/ → [b]	/b/ → [b], [b ^h]	/B/ → [b], [...], [b ^h]
/b ^h / → [b ^h], [b]		+voiced, ?aspirated]
/b ^h au/	/bau/	/Bau/

The problem with the first alternative (i.e. postulating two phonemes /b/, /b^h/), is that it means positing a phoneme that is only occasionally contrastive on the surface and often has the exact same realization as another phoneme. Additionally, this merger is not rule-bound but completely free. The problem with the second alternative is that it brings a very strange kind of allophony with it. There is free variation, but this free variation occurs only in certain lexical items. Something being lexically restricted is generally taken to be a criterion for phonemicity. The third alternative sneaks in lexical restriction through an underspecified underlying representation of unclear status vis-à-vis the empirical language data.

Finally, we may note one potential minimal pair in our materials: [g^hoɾa] ‘horse’ and [goɾa] ‘ankle’. There are, however, two problems with treating this minimal pair as conclusive evidence for phonemic status. The first is the already mentioned free variation: ‘horse’ is just as often (probably more often, though it is hard to quantify precisely from our data) realized as [goɾa], homophonous with ‘ankle’.

4.2 Phoneme inventory: geminate consonants

We analyze Kanashi as having phonemic geminates for all stops and fricatives and for /l/. Our material contains no minimal pairs, but pairs such as /batak/ ‘duck’ and /battis/ ‘thirty-two’ demonstrate the difference. Instrumental measurements show a difference in duration between geminates and singletons (Saxena et al. 2018). See Table 3 for some examples.⁵

Table 3: Geminate consonants in Kanashi

Kanashi (YF)	Gloss
/didd/	‘there’
/ɖikke/	‘anger’
/massi/	‘mother’s sister’
/tʰulla/	‘leg, foot’

As with the aspirated voiced stops, there are, however, some issues complicating the incorporation of geminates into the phonological analysis of Kanashi.

The first is that there appears to be a clear difference between speakers. In the speech of YM, consonants that in the speech of YF are heard as geminate are not heard as geminate and show no significant durational difference in instrumental measurements (Saxena et al. 2018); see Figures 1 and 2.

Figure 1: geminate (left: /kutta:/ ‘dog’) : singleton (right: /kuta:b/ ‘book’) (YF)

⁵ Approximately 110 items are found in our material that contain geminates. The number varies, depending on how we count items that appear both with and without a geminate, e.g. /suk(k)ar/ ‘Friday’.

Figure 2: geminate (left: /kutti/ ‘bitch’) : singleton (right: /kʰutu/ ‘penis’) (YM)

Thus, the presence of geminates appears not to be constant across speakers, which brings up the question which speakers’ variety should be taken as the basis for the phonology. We see some intra-speaker variation in our data, but not to the same extent as when it comes to voiced aspirates. Some examples are given in Table 4.

Table 4: Variation in consonant gemination

YF consultant	YM consultant	Gloss
[daddu]	[dadu]	‘paternal grandfather’
[dubbem]	[dubem]	‘to be drowned’
[matʰa]	[matʰa]	‘forehead’
[haɖɖaŋ]	[haɖaŋ]	‘bone’
[duppe]	[dup(p)e]	‘sun’

Secondly, just as in the case of the aspirated voiced stops, geminates appear to be found mainly in IA items, so that the observed variation may simply be due to the more educated speakers showing less integration of IA items in their Kanashi.

5 Discussion

As we have seen, several factors appear to play a role in the phonetic variation found in Kanashi. The details of their interplay and which factors are responsible for what is as of yet not entirely clear.

One of these factors is sociolinguistic aspects. As mentioned, we find regular differences between our consultants, such as with the geminates, described

above. It is of course tempting to draw sociolinguistic conclusions from this, such as connecting the literacy of YF with the features she exhibits. However, this is risky to do without first collecting more data from other speakers, representing different social variables. This, then, is one of the ways forward to gaining a fuller picture of variation in Kanashi.

Another factor is the distribution of these features across different language strata. In some languages with lexical items of different origin, it may be motivated with different phonologies for different layers in the languages. An example of where such an analysis has been suggested is Japanese, see e.g. Ito & Mester (1995). We have already hinted at why such an analysis seems less desirable for Kanashi. As already mentioned, none of the features discussed are strictly limited to words of IA origin. This sets these features apart from, for instance, the cluster /ks/, which is very clearly limited to a handful of IA/English loans.

It should also be noted that we occasionally find one of these features in items where they do not exist in Kullu Pahari or Hindi. It is currently not clear if this is due to internal change in Kanashi (for instance, by analogy with other borrowed items) or if the items were borrowed in that form.

A final factor to be considered is the possibility that even with comprehensive data, the status of some features in Kanashi will remain unclear, or at least less clearly phonemic than some others. Scobbie & Stuart-Smith (2008: 106) write, regarding some problematic features of Scottish Standard English, that “[i]t must not be thought that these difficulties arise due to sociolinguistic or stylistic variation, and that they can be dismissed as just so much ‘noise’ by researchers whose focus is exclusively phonological theory. We think that any variation presented above is relevant to phonology in the narrowest sense”. They suggest treating phonemicity as a graded property. An overview of similar problems and suggested solutions can be found in Hall (2013), where she refers to them as “intermediate phonological relationships”. That is, relationships that are not clearly either phonemic or allophonic but seem to lie somewhere in between, either because the traditional criteria for phonemic status conflict or because some other factors are felt to affect their status. The notion is intuitively appealing, but as usual the devil is in the details. Gradient phonemic status requires a way of calculating how much of a phoneme a particular segment is, or the notion risks becoming vacuous. While Hall (2013: 259–262) discusses at some length various proposals for calculating degree of phonemicity, no immediately practical procedure for doing this is presented. In the absence of such, observations about phoneme gradience or marginal phonemes may be enlightening, but with limited utility, and in reality the only reasonable practical descriptive solution may be a binary phoneme ~ allophone distinction.

We will here point out some ways in which the Kanashi features discussed here make a classical binary divide into phonemic and allophonic status problematic.

With regard to the variations observed concerning geminates and aspirated voiced stops, our preliminary analysis is as follows: aspirated voiced stops cannot be considered fully part of the phoneme inventory of Kanashi; their rarity, their limited distribution strata-wise and the fact that they are always in free variation with plain voiced stops make this clear. Of these, the last one is the strongest argument. However, they also cannot be considered merely allophones of their plain voiced stop counterparts either: the main piece of evidence for this is that they are lexically conditioned, and not phonologically conditioned or in completely free variation. They seem to be examples of so-called marginal phonemes, a time-honored, seemingly unavoidable concept in actual language descriptions (e.g. Ferguson & Chowdhury 1960; Suomi et al. 2008). Returning to the three alternatives discussed above in Section 4.1 (repeated here for convenience), we note the following.

Alternative 1	Alternative 2	Alternative 3
/b/ → [b]	/b/ → [b], [b ^h]	/B/ → [b], [...], [b ^h]
/b ^h / → [b ^h], [b]		+voiced, ?aspirated]
/b ^h au/	/bau/	/Bau/

The amount of information that needs to be captured in any language description is determined by attested language phenomena and equal regardless of the particular linguistic model adopted.⁶ That said, there are better and worse models, of course. Model-internal parsimony – Occam’s razor – is a desirable goal. But in principle there are many different models which will capture the same information equally economically, and the choice among these must be made according to some other criterion or criteria, e.g. relating aspects of the linguistic model to findings from neuro- or psycholinguistics. The main point we wish to make here is simply that the attested facts of the (lexically conditioned) distribution of aspirated and unaspirated voiced stops in Kanashi must be accounted for in our description, and there are several ways in which this can be done.

⁶ “Information” is understood here in the technical sense of mathematical information theory (Shannon & Weaver 1949), which defines a lower bound on the number of symbols needed to faithfully express a certain set of distinctions.

Our general preference – at least at this stage of description – is to make as few assumptions as possible necessitating the postulation of underlying, ‘invisible’ entities. Hence, morphemes should be realized as (one or more) phoneme sequences without underspecified segments (such as archiphonemes). Phonemes should in principle correspond to one of their actually occurring allophones. This rules out alternative 3. Further, allophone selection should preferably be determined only by phonological (or phonetic context). This rules out alternative 2, since, as mentioned above, the allophone selection must be lexically determined in our case. The remaining option, alternative 1, is in some sense the mirror image of alternative 2. However, if there are any nonalternating instances of aspirated voiced stops, this alternative will overgenerate, unless we again introduce the possibility of lexical triggering of a phoneme realization rule.⁷

A relevant consideration, and one that Hall (2013) does not mention as such, although it is implicit in some of her argumentation, has to do with what we could call “model-internal consistency”. Some modelling assumptions logically restrict other aspects of our model. In fact, if we take the segmental representation(s) of a morpheme to be made up of phonemes, as above, this logically means that any allophony must be computable from this representation (since the representation itself is made up from phonemes). If we further require the conditioning factors for selecting allophones to be strictly phonological, we are forced to accord the aspirated voiced stops phoneme status in Kanashi, and assume that the relevant morphemes have two allomorphs listed, whose precise selection criteria remain to be elucidated, however.

Against this background, the safest route making the fewest assumptions would be to assume that a lexical item like ‘brother’ will be realized by two possible phoneme sequences: /b^hau/, /bau/. Note that alternative 1 and this solution both require that aspirated voiced stops be recognized as phonemes (although marginal).

As for geminates, we have pointed out that the difference in realization here seems mainly to be between speakers. If this turns out to hold, it may be that we can in the future define geminates as sociolinguistically determined. Whether geminates are to be considered a part of the Kanashi sound system as a whole is another question, and must presumably depend on how many and which speakers have it.

⁷ Since we cannot in principle demonstrate non-occurrence of variation in an individual case, this alternative still seems as the preferable one.

6 Conclusions

We have shown in this chapter that several factors, many of which are perhaps not foremost in the researcher's mind when doing early-stage fieldwork, can be relevant in describing the variation and phonology of a language. Such variation poses additional challenges when it comes to describing the language (sound system) for the first time for a language where the available language data as well as the access to native speakers are very limited. The examples from Kanashi show that we must take into account both potential sociolinguistic factors as well as the etymology of lexical items. Additionally, it should be considered that the conventional criteria for phonemic status might not give a clear answer, regardless of how thorough the coverage of the materials is of the phenomenon under investigation.

References

- Andersen, Roger W. 1982. Determining the linguistic attributes of language attrition. In Richard D. Lambert & Barbara F. Freed (eds.), *The loss of language skills*, 83–118. Rowley: Newbury House.
- Boersma, Paul & David Weenink. 2018. *Praat: Doing phonetics by computer [computer program]*. Version 6.0.42, retrieved 15 August 2018 from <http://www.praat.org/>.
- Cook, Eung-Do. 1989. Is phonology going haywire in dying languages? Phonological variations in Chipewyan and Sarcee. *Language in Society* 18(2): 235–255.
- Darnell, Regna. 1990. *Edward Sapir: Linguist, anthropologist, humanist*. Berkeley: University of California Press.
- Ferguson, Charles A. & Munier Chowdhury. 1960. The phonemes of Bengali. *Language* 36(1): 22–59. DOI: 10.2307/410622.
- Hall, Kathleen C. 2013. A typology of intermediate phonological relationships. *The Linguistic Review* 30(2): 215–275.
- Ito, Junko & R. Armin Mester. 1995. Japanese phonology. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 817–838. Oxford: Blackwell.
- Joseph, John Earl. 1987. *Eloquence and power: The rise of language standards and standard languages*. London: Pinter.
- Linell, Per. 2005. *The written language bias in linguistics*. First published in 1982 by Dept. of Communication Studies, University of Linköping, Sweden. London: Routledge.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. Berkeley: University of California Press.
- McGregor, R. S. 1993. *The Oxford Hindi-English dictionary*. Oxford: Oxford University Press.
- Palosaari, Naomi & Lyle Campbell. 2011. Structural aspects of language endangerment. In Peter Austin & Julia Sallabank (eds.), *Cambridge handbook of endangered languages*, 100–119. Cambridge: Cambridge University Press.

- Saxena, Anju, Anna Sjöberg & Padam Sagar. 2018. *Linguistic variation: Challenges for describing the phonology of Kanashi*. Presentation at the *International workshop on Fieldwork: Methods and theory*. University of Gothenburg, Sweden, 13–14 December 2018.
- Scobbie, James M. & Jane Stuart-Smith. 2008. Quasi-phonemic contrast and the indeterminacy of the segmental inventory: Examples from Scottish English. In Peter Avery, B. Elan Dresher & Keren Rice (eds.), *Contrast in phonology: Perception and acquisition*, 87–113. Berlin: Mouton de Gruyter.
- Shannon, Claude E. & Warren Weaver. 1949. *The mathematical theory of communication*. Champaign: University of Illinois Press.
- Suomi, Kari, Juhani Toivanen & Riikka Ylitalo. 2008. *Finnish sound structure: Phonetics, phonology, phonotactics and prosody*. Oulu: Oulu University Press.