# Numerals in Tibetan

## Jeffrey Heinz and Kristina Strother-Garcia

## February 21, 2016

This chapter provides a computational analysis of the variation in the pronunciation of some Tibetan numbers.[1] In Tibetan, the names for numbers greater than ten are formed by concatenating the word for a number less than ten with the word that means 'ten.' We refer to these combinations as "compound numerals." In particular, numbers from eleven to nineteen are formed by prefixing [d͡ʒu] 'ten' to the number being added to ten. For example, [d͡ʒuŋa] 'fifteen' has the number [ŋa] 'five' as the first element of the compound, so it can be analyzed as [d͡ʒu-ŋa], literally meaning 'ten-five.' To form multiples of ten greater than ten, the root meaning 'ten' is concatenated as a suffix to the number being multiplied by ten, as in [ŋabd͡ʒu] 'fifty' (literally 'five-ten'). This example shows that there is some variation in pronunciation across different compound numerals. While it appears that [ŋa] 'five' is prefixed to [d͡ʒu] 'ten', we also see the additional segment [b] between the two. As the table below shows, there are a number of other such consonant alternations (highlighted in bold).

| | | | | | |
|---|---|---|---|---|---|
| d͡ʒig | 'one' | d͡ʒug**d͡ʒ**ig | 'eleven' | | |
| ʃi | 'four' | d͡ʒu**b**ʃi | 'fourteen' | ʃibd͡ʒu | 'forty' |
| ŋa | 'five' | d͡ʒuŋa | 'fifteen' | gubd͡ʒu | 'ninety |
| gu | 'nine' | d͡ʒu**r**gu | 'nineteen' | ŋa**b**d͡ʒu | 'fifty' |
| d͡ʒu | 'ten' | | | | |

Table 1: Tibetan numbers and compound numerals.

We see that [b] follows the first root and precedes [d͡ʒu] in all multiples of ten greater than ten: [ʃibd͡ʒu] 'forty,' [ŋabd͡ʒu] 'fifty,' and [gubd͡ʒu] 'ninety.' In each case, [b] is preceded by a vowel and followed by the affricate [d͡ʒ]—there is no more specific generalization we can make about the environment of this [b]. Assuming that this [b] does not appear in the underlying form of either the root meaning 'five' or the root meaning 'ten,' we might propose some phonological process that epenthesizes [b] between a vowel and the affricate [d͡ʒ]. However, this process would incorrectly predict the surface form [*d͡ʒubd͡ʒig] instead of [d͡ʒugd͡ʒig] 'eleven.'

Furthermore, as Table 1 shows, the roots used to form other numbers in the teens vary in pronunciation as well. 'Fourteen' is [d͡ʒubʃi] rather than [*d͡ʒuʃi] and 'nineteen' is [d͡ʒurgu] rather than [*d͡ʒugu]. This variation is not easily explained by a single process of epenthesis because the consonants that appear in the compound numerals are not predictable based on the environments. Although we may be able to conceive of multiple epenthetic processes that could explain all of this variation, there is a much simpler answer: the underlying forms of certain numerals contain a consonant which is absent in the morphologically bare surface form.

Under this latter hypothesis, there are several possible combinations of underlying forms one might suggest. Let us first reconsider the example above: [ŋabd͡ʒu] 'fifty' can either be morphologically analyzed as /ŋab-d͡ʒu/ or /ŋa-bd͡ʒu/. If the underlying form of 'ten' were /d͡ʒu/ and the underlying form of 'five' were /ŋab/, we might suppose that the underlying forms of 'four' and 'nine' were /ʃib/ and /gub/, respectively, conforming to the pattern we see in multiples of ten. We could explain the surface forms of the multiples of ten by positing a phonological process which deletes word-final consonants. This would accurately predict the absence of a word-final [b] in the morphologically bare surface forms and in the surface forms for 'fourteen', 'fifteen', and 'nineteen'; it would also correctly predict the word-internal [b] that appears in the surface forms for 'forty', 'fifty', and 'ninety'. However, the surface forms [d͡ʒig] 'one' and [d͡ʒugd͡ʒig] 'eleven' would not be predicted because both have a word-final consonant, [g]. We could modify our hypothesis to say that only /b/ is deleted word-finally, but such a specific process seems ad hoc. It also fails to explain why the surface forms [d͡ʒubʃi] 'fourteen' and [d͡ʒurgu] 'nineteen' also include consonants not present in the underlying forms of the supposed roots that make up these compound numerals.

A more reasonable alternative is that the underlying form of 'five' is actually [/ŋa/] and the underlying form of 'ten' is /bd͡ʒu/. Under this assumption, we would also posit the underlying forms: /gd͡ʒig/ 'one,' /bʃi/ 'four,' and /rgu/ 'nine.' The surface variation among numerals and their roots in compound numerals can be explained by a phonological process which deletes the first consonant of a word-initial consonant cluster (CC). Hence /bʃi/ surfaces as [ʃi] 'four.' This would accurately predict the absence of word-initial consonant clusters throughout the data, as well as the pattern of consonants seen in compound numerals. Because this explanation is both accurate and economical (in the sense that it only requires one phonological process to predict all observed variation in surface forms), this hypothesis is preferred over the ones previously mentioned.

The theory of prosodic licensing (Ito, 1986; Itô, 1989) provides a reason why initial consonant clusters may be simplified. Under this theory, extra-syllabic segments fail to surface. In other words, segments which fail to be incorporated into syllables are deleted, a process known as *stray erasure* since the "stray" consonants are erased (McCarthy, 1979; Steriade, 1982; Harris, 1983).

In what follows, we provide a computational analysis of this generalizations. Underlying representations are analyzed in terms of a word model which contains the binary successor relation (◁) for order with a standard set of phonological features $\mathcal{F}$ represented as unary relations. In this analysis we will refer specifically to the features `consonantal` and `vocalic`.

Underlying word model

$$\langle \mathcal{D}, \lhd, \mathtt{feature}\ (\forall\ \mathtt{feature} \in \mathcal{F}) \rangle$$

Surface word model

$$\langle \mathcal{D}, \lhd, \mathtt{onset}, \mathtt{coda}, \mathtt{nucleus}, \mathtt{feature}(\forall\ \mathtt{feature} \in \mathcal{F}) \rangle$$

Figure 1: Models for underlying and surface forms.

We will not specify other features in $\mathcal{F}$ but we assume they are there and are sufficient to distinguish the transcribed speech sounds in Tibetan.

Per standard phonological theory, syllabic structure is only present in surface forms, and not in underlying forms. Therefore, the surface representations are analyzed with a word model that includes —like the model for underlying forms—the binary successor relation ($\lhd$) for order and a standard set of phonological features $\mathcal{F}$, but also includes—unlike the model for underlying forms—unary relations describing syllabic roles such as $\mathtt{onset}$, $\mathtt{coda}$, and $\mathtt{nucleus}$.

These models are summarized in Figure 1.

With these models in place, we can begin to specify the transformation from underlying to surface forms. Recall that in order to specify this transformation, we must specify the following items.

- A formula with no free variables that establishes the domain of the transformation $\varphi_{dom}$. This determines those structures to which the function can apply.

- A copy set $C$ of $k \geq 1$ elements which determines, along with the size of the input's domain, the maximal size of the output's domain. Each pairing $(c, x)$ with $c \in C$ and $x$ in the domain of the input form is a possible element in the domain of the model of the surface form.

- For each element $c$ in the copy set, a licensing formula of one free variable $\varphi_{license}^c(x)$ which determines whether $(c, x)$, which is the $c$th copy of element $x$, is licensed in the model of the surface form. Unlicensed elements are not part of the domain of the surface form, but licensed ones are.

- For each binary relation in the model of the surface forms, a formulae of two free variables. Here, there is only one such formula $\varphi_{\lhd}$.

- For each unary relation in the model of the surface forms, a formula of one free variable. Hence, we will have one for each feature in $\mathcal{F}$ as well as one for each of $\mathtt{onset}$, $\mathtt{coda}$, and $\mathtt{nucleus}$.

Also, the formulae can be specified in any order as long as they are well-defined. Thus one formula $\varphi$ can be defined in terms of another formula $\psi$ only if $\psi$ has been defined previously.

We now define a transformation from underlying to surface forms which captures the aforementioned linguistic generalizations in Tibetan. Specifically, we will enforce deletion of unsyllabified consonants through the licensing formula.

We assume the domain of the transformation is any structure which can be represented by the underlying word model.

$$\varphi_{dom} \overset{\text{def}}{=} \texttt{true} \tag{1}$$

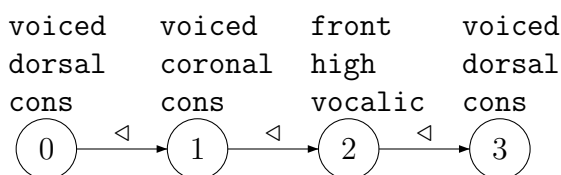For example, a valid input is given by /gd͡ʒig/ 'one' is shown in Figure 2.



Figure 2: A graph representing the word model of /gd͡ʒig/ 'one.' If $x \lhd y$ then there is a labeled directed edge from $x$ to $y$. Some unary relations to which element $x$ belongs are shown above $x$.

Since the size of the domain of every output is no larger than the size of the input we set the copy set to contain a single element.

$$C \overset{\text{def}}{=} \{1\} \tag{2}$$

Consequently this means the size of the domain of the output is maximally the size of the input domain. This "workspace" is illustrated in Figure 3.



Figure 3: The possible elements of the domain of the model of the surface form.

Next, Equation 3 defines the the binary successor relation in the output form.

$$\varphi_{\lhd}(x, y) \overset{\text{def}}{=} x \lhd y \tag{3}$$

$\varphi_{\lhd}(x, y)$ will evaluate to true only if it is true that $x \lhd y$ in the model of the underlying form. If $\varphi_{\lhd}(x, y)$ is true then $(x, y)$ will stand in the successor relation in the model of the surface form. Figure 4 shows which relations would be present if every possible domain element in the workspace is licensed.

Next, for each unary relation in the model for the surface forms, we need a formula which identifies those elements that property. Since no element changes features, this is relatively straightforward.

$$(\forall \texttt{ feature} \in \mathcal{F}) \ \ \varphi_{\texttt{feature}}(x) \overset{\text{def}}{=} \texttt{feature}(x) \tag{4}$$
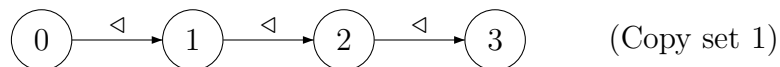
Figure 4: The binary relations that would be present in the model of the surface form if every domain element is licensed.

Equation 4 is actually several equations that are all similar in character, which is why we bundle them together. Essentially it says that if an element $x$ in the underlying model belonged to a unary relation `feature` then its corresponding element $x$ also belongs to that unary relation. In other words, elements in the surface form (if they exist) carry the same features they carried in the underlying form.

Of more interest is of course the unary relations which determine the syllabic roles syllables will play. These are defined next.

$$\varphi_{\texttt{nucleus}}(x) \quad \overset{\text{def}}{=} \quad \texttt{vocalic}(x) \tag{5}$$

$$\varphi_{\texttt{onset}}(x) \quad \overset{\text{def}}{=} \quad \texttt{consonantal}(x) \land (\exists y)[\texttt{vocalic}(y) \land x \lhd y] \tag{6}$$

$$\varphi_{\texttt{coda}}(x) \quad \overset{\text{def}}{=} \quad \neg\varphi_{\texttt{onset}}(x) \land \texttt{consonantal}(x) \land (\exists y)[\texttt{vocalic}(y) \land y \lhd x] \tag{7}$$

Essentially, what these equations guarantee is that vowels will be nuclei, consonants immediately preceding vowels will be onsets, and non-prevocalic consonants immediately succeeding vowels will be codas. Consequently, as illustrated in Figure 5, if every possible element in the model of the surface form is realized, the elements would belong to the unary relations as shown.
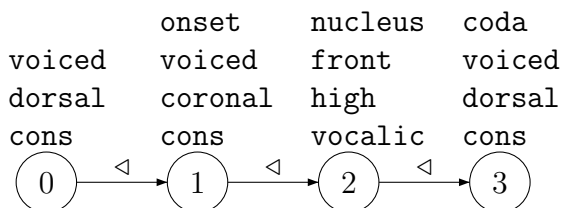


Figure 5: The binary and unary relations that would be present in the model of the surface form if every domain element is licensed.

The last formula needed to specify the transduction is the formula which indicates which of the possible elements in the workspace are actual elements. In our case, this is simply those elements that have syllabic roles.

$$\varphi_{\texttt{licensed}}(x) \overset{\text{def}}{=} \varphi_{\texttt{nucleus}}(x) \lor \varphi_{\texttt{onset}}(x) \lor \varphi_{\texttt{coda}}(x) \tag{8}$$

Elements in the domain which are licensed are precisely those which belong to one of the unary relations. Clearly, every element $x$ except for $x = 0$ makes $\varphi_{\texttt{licensed}}(x)$ true. Hence, in the running example, element 0 is not part of the final model.

Together, Equations 1 to 8 define a transduction that transforms models of underlying forms to models of surface forms. There is a domain formula ($\varphi_{dom}$), a copyset $C$, one formula of two free variables ($\varphi_\lhd(x, y)$), formulae of one free variable for each unary relation in the model of the surface forms, and finally one formula of one free variable which determines which of the possible elements are actually realized.

Figure 6 illustrates the model output by the specified transformation when given the model of the input [d͡ʒig] 'one,' shown in Figure 2.

```
onset      nucleus   coda
voiced     front     voiced
coronal    high      dorsal
cons       vocalic   cons
 (1)   ◁    (2)   ◁    (3)
```
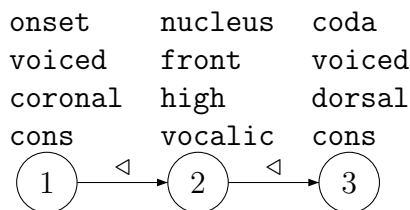
Figure 6: A graphical visualization of the relational structure which is the output of the input [d͡ʒig] 'one' shown in Figure 2 under the transductions defined in Equations 1 through 8.

The prosodic theory of licensing is of course not the only way to analyze the alternation in Tibetan numerals. Another analysis could have forgone the syllabic roles and simply licensed all positions except word-initial consonants immediately succeeded by other consonants. Such an analysis would also successfully account for the present data. These two analyses however are not extensionally equivalent: the prosodic licensing analysis predicts that if triple consonant clusters were to occur word-internally in underlying forms, the middle consonant would delete (since it would be the only one unsyllabified). However the latter hypothesis makes no such prediction.

Another analysis is possible in Optimality Theory (Prince and Smolensky, 2004). Here, similar to the prosodic licensing analysis, syllable structure plays a motivating role. An unbounded markedness constraint banning complex syllable margins motivates consonant cluster reduction. The first consonant, and not the second, deletes because of the faithfulness constraint IO-CONTIG which prefers deletions to occur at edges and not internally. Unlike the prosodic licensing analysis however, this analysis as stated makes no predictions which consonant deletes in word-internal triple consonant clusters. It would fall to other constraints to determine which one deletes (for instance the least marked consonant, whichever one it may be).

In the second part of this chapter, we explore how the generalizations identified in the aforementioned OT analysis may be presented in the computational framework utilized here.

[To be completed...]

# References

Halle, Morris, and G. N. Clements. 1983. *Problem Book in Phonology*. Cambridge, MA: MIT Press.

Harris, James. 1983. *Syllable Structure And Stress in Spanish: a Nonlinear Analysis*. Cambridge, Mass.: MIT Press.

Ito, Junko. 1986. Syllable Theory in Prosodic Phonology. Doctoral dissertation, University of Massachusetts, Amherst. Published 1988. Outstanding Dissertations in Linguistics series. New York: Garland.

Itô, Junko. 1989. A prosodic theory of epenthesis. *Natural Language & Linguistic Theory* 7:217–259.

McCarthy, John J. 1979. On stress and syllabification. *Linguistic Inquiry* 10:443–466.

Odden, David. 2014. *Introducing Phonology*. 2nd ed. Cambridge University Press.

Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell Publishing.

Steriade, Donca. 1982. Greek prosodies and the nature of syllabification. Doctoral dissertation, MIT, Cambridge, Mass.