### Standard Minangkabau | Sorba | English Translation
---|---|---
a | raso | sora | 'taste, feeling'
b | rokok | koro | 'cigarette'
c | rayo | yora | 'celebrate'
d | susu | sursu | 'milk'
e | baso | sorba | 'language'
f | lamo | morla | 'long time'
g | mati | tirma | 'dead'
h | bulan | larbu | 'month'
i | minum | nurmi | 'drink'
j | lilin | lirli | 'wax, candle'
k | mintak | tarmin | 'request'
l | cubadak | darcuba | 'jackfruit'
m | mangecek | cermange | 'talk'
n | bakilek | lerbaki | 'lightning'
o | sawah | warsa | 'rice field'
p | pitih | tirpi | 'money'
q | manangih | ngirmana | 'cry'
r | urang | raru | 'person'
s | apa | para | 'father'
t | iko | kori | 'this'
u | gata-gata | targa-targa | 'flirtatious'
v | maha-maha | harma-harma | 'expensive'
w | campua | purcam | 'mix'
A comparison of a-c would indicate that to form a Sorba word one takes the consonant and vowel of the last syllable, e.g. so from raso, ko from rokok and yo from rayo and one places it at the beginning of the word. If the last syllable ends in a consonant, e.g. final k in rokok then one deletes it.

So we might state the rules as:
1. Delete the word final consonant: (rokok > roko)
2. Take the final syllable (or C + V) and make it the first syllable (roko > koro, raso > soro, rayo > yora)

However, if we apply these rules to the following words (d-w) we fail to create the correct Sorba word. We notice that a common feature of Sorba words is that the third sound MUST BE r. So we need a rule which inserts r, unless the standard language word begins with r. Notice how this is requirement for a Sorba word.

As we need to stipulate that the third sound must be r, we must add another rule:
Rule 3: Add r to initial CV unless the following sound is r.

Notice that we have to spell out the condition in which the rule applies (i.e. in the absence of following r), so we don't get a sequence of r+r.

We can see from examples h-r that our rule 1 applies.

The reduplicated words in u and v show us that each part of the reduplication must undergo the Sorba formation rules, e.g., gata-gata > targa-targa (NOT targataga). So we would need to stipulate that reduplicated words are treated like two words, and not as a single word.

Example w campua > purcam (NOT paurcam or arcampu). This shows us that only the initial Consonant and Vowel of the final syllable is moved to the front of the word to form a Sorba word, so that we need to modify our Rule 1. Furthermore, a word final vowel which follows another vowel is not treated as a final syllable for the Sorba formation.

Modified Rule 1: Delete any sound which follows the final CV sequence.

So after inspection of all the words we can express the rules for converting a standard Minangkabau word into a Sorba word as:

Rule 1. Treat reduplicated words as a sequence of two identical words.
Rule 2. Delete any sound which follows the final CV sequence of a word.
Rule 3. Move the final CV sequence to the start of the word
Rule 4. If the third sound of the new word is not r, insert r (after the first CV sequence).
H1.

<table>
<thead>
<tr>
<th>Standard Minangkabau</th>
<th>Sorba</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>rancak</td>
<td>caran</td>
<td>'nice'</td>
</tr>
<tr>
<td>jadi</td>
<td>dirja</td>
<td>'happen'</td>
</tr>
<tr>
<td>makan</td>
<td>karma</td>
<td>'eat'</td>
</tr>
<tr>
<td>marokok</td>
<td>kormaro</td>
<td>'smoking'</td>
</tr>
<tr>
<td>ampek</td>
<td>peram</td>
<td>'hundred'</td>
</tr>
<tr>
<td>limpik-limpik</td>
<td>pirlim-pirlim</td>
<td>'stuck together'</td>
</tr>
<tr>
<td>dapua</td>
<td>purda</td>
<td>'kitchen'</td>
</tr>
</tbody>
</table>

H2. We can only work back to a set of possible standard Minangkabau words because of two difficulties or problems:

'r' problem: we can't know if 'r' in lore was in standard word or whether it was inserted by Sorba 'r' insertion rule, e.g., standard elo or relo > Sorba lore

final sound problem: we can't know if standard word ends in consonant or one or two vowels or not as Sorba deletes final consonant/vowel following a vowel. lore could be derived from elo, relo, eloa, reloa or eloC or reloC where 'C' stands for any possible final consonant.

H3. We can see that the word formation rules for converting a Minangkabau word into Solabar are:

1. Delete the sound which follows the final CV sequence.
2. Move the final CV sequence to the beginning of the word.
3. Add la to the new word initial CV sequence
4. Delete the sound which follows the new final CV sequence
5. Add r to the word.

In converting baso we don't need to apply Rule 1.
We apply rule 2 > soba
We apply rule 3 > solaba
We don't need to apply rule 4.
We apply rule 5 > solabar

In converting campua and makan:
Rule 1 applies: > campu > maka
Rule 2 applies: > pucam > kama
Rule 3 applies: > pulacam > kalama
Rule 4 applies: > pulaca (doesn't apply)
Rule 5 applies: > pulacar > kalamar
The Solabar equivalent of the Sorba word *tirpi* 'money' is *tilapir*.

To answer this question we need to reconstruct the form of the Standard word. Luckily for us it is given in the initial list (p.) as *pitih*. By applying our rules we get: *pitih > piti > tipi > tilapi > tilapir*.

However, if instead of assuming that rule 3. is "add la..." which we cannot be sure about from the data given, since the syllable following *la* has the vowel 'a' in all three words (plus *solabar*), it is possible that the rule should be add l+vowel where vowel is a copy of the following vowel. This would then open the possibility that our answer could be *tilipir*. Now given that the final vowel of our Solabar data set only contains the vowel *a*, maybe our rules 4 and 5 should really be collapsed to a single rule: "substitute *ar* for the sound or sounds which follow the first C in the 'new' final syllable. If we applied this rule and allowed for the other two possibilities we would have to allow the possibility of getting Solabar forms: *tilipir, tilapir, tilipar, tilapar*.

In order to disconfirm the incorrect hypotheses we would need to see how a Minangkabau word such as *lilin* 'wax' forms its Solabar form. If it is *lilalir* then we know that our original rules are correct. If it is *lilalar* we know that we need to change our rules. Notice that if the final syllable were always required to end in *ar* then there is no way of distinguishing between our Rule 3 "Add la..." or a rule which says "Add lV where V = same as V in final syllable".

Notice that a rule which requires the final syllable to end in *ar* would make for a more complex set of rules. As our rules stand, Rules 1 and 4 are identical - they just apply at different stages in the word formation process. This would not be the case if the Solabar words had to end in *ar*; Rule 4 would be different from Rule 1.

**H4.** 'ng' is one sound because the Sorba for standard Minangkabau *manangih* 'cry' is *ngirmana*. if 'ng' were two sounds the Sorba word would begin with *g* and end in *n* by our rules. i.e., *girmanan*.

Notice that we would need to create some special specific complicated rules to get a sequence of two consonants (as opposed to two letters representing a single sound) at the beginning of this Sorba word and to exclude them for other words, e.g., how would we prevent *mintak* from being coverted to Sorba *ntarmi* rather than the correct *tarmin*?

We are always looking for the simplest solution or explanation to account for the facts we observe.