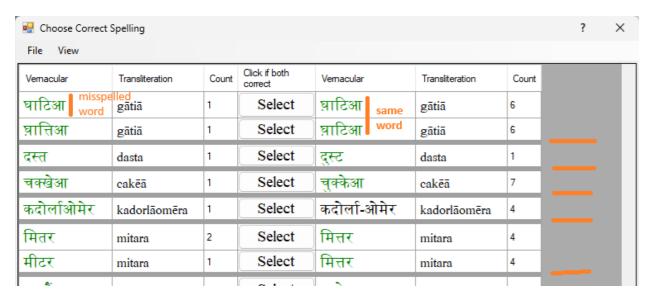
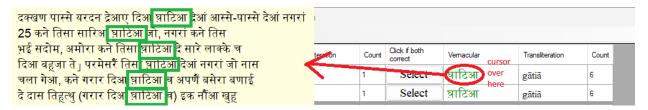
Cheat sheet for Consistent SpellFixer

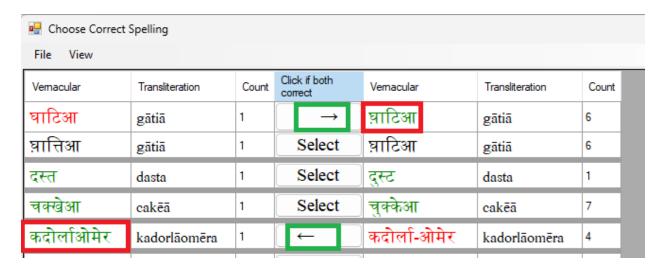
- Basic Procedure:
 - 1. Open the Bulk SFM Converter
 - 2. Click the File, Open SFM Documents, Unicode
 - 3. Select all the SFM Documents you have in your corpus (e.g. all the SFM documents in, say, C:\My Paratext 9 Projects\project> folder). You might not want to include any that don't have words you want to check the spelling of.
 - 4. Click on the Consistent Spelling Options menu, and the Select Consistent Changes Project... menu. If there are no projects shown, then click the Add New Project button.
 - 5. Fill in the information (contact <u>silconverters_support@sil.org</u> for help depending on the language family you are using, but often the defaults can be used).
 - 6. Once you select the project and click *OK*, the converter for the SFM markers commonly having data that needs to be spell checked will be initialized with the name of the converter used for fixing bad spellings (e.g. *Consistent Spelling for Kangri*) so that the data in those markers will be checked for similar spelling. You can click on the center column buttons if you want to include other fields that aren't initialized (they'll say "Click to check spelling" if they're not initialized).
 - 7. Click on the Consistent Spelling Options menu again, and the Analyze Data for Consistent Spelling menu. This will launch another dialog (see examples below) that will have similarly spelled words in one or more rows in bundles of ambiguous words. This dialog loads in the background, but you don't have to wait for it to finish before beginning the selecting process discussed below:
- Some rows of the table have bigger spaces above and below. All the rows inside the bigger space boundaries are "ambiguous" with each other. It might be a single row of a left and right word that are ambiguous with each other, or it might be multiple rows where the right-hand side word of all rows in the 'bundle' is the same word and is the most frequently used spelling (and therefore more likely might be the correct spelling of one or more of the words on the left). The left-hand side words are the various other words that are part of the same ambiguous bundle:



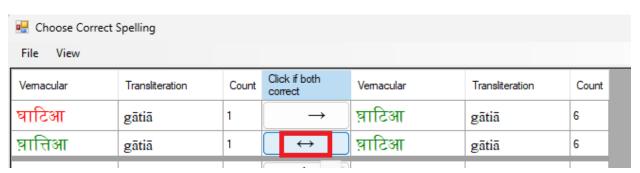
• Hover your cursor over the words in either the left or right-hand side *Vernacular* column to see a yellow popup window in upper left corner of screen showing some of the contexts in which the word appears:



- The Count column value indicates relative frequency of words (up to 20, but this is configurable in the project file, search for the MaxContextStringsInToolTip attribute in the projectName.csf file, located in by default in the C:\ProgramData\SIL\ConsistentSpellingFixer\Projects\projectName folder).
- Left click on the words in the Vernacular column to indicate that either the left or
 right side word is the correct spelling of the other. This will put an arrow pointing
 to the correctly spelled word in the center column and make the good spelling
 show as green color and the bad spelling word as red:



• If both words are correct, then left-click (once) on the word "Select" in the center column and you'll get a double-headed arrow:



- You can left-click again on the center button to undo a selection made previously.
- If a word (usually on the left) is bad, but the word on the right is not the correction of it, you can do one of two things:
 - 1. If it's a multi-line bundle of ambiguities, you can right-click on the correctly spelled word in the left-hand side (of some bad spelled word somewhere else either on the left or right-hand side) and the word you click on will become the new right-hand side word for each row. This allows you to then click on the new correctly spelled word on the right-hand side that's the correction of one or more of the words on the left-hand side.

For example, suppose that before right-clicking on the left-hand side word, the situationw as this:

П			•					
	कहड्डे go	ood	kahadē	1	Select	खहुड्डे	kahadē	8
	कहुडे b	ad	kahadē	1	Select	खहुड्डे	kahadē	8

One of these left-hand words is the correction of the other, while the right-hand word has no relationship to them. So you could right-click on the correctly spelled word in the left and that word will go to the right-hand side

After right-clicking on the good word on the left:

खहुड्डे	kahadē	8	Select	कहड्डे g	good	kahadē	1
कहुडे	kahadē	1	Select	कहुड्डे	word	kahadē	1

After the good word moves to the right-hand side, you can click on it in the 2nd row to show that it's the correct spelling of the left-hand side word in the 2nd row

खहुड्डे	kahadē	8	Select	कहड्डे		kahadē	1
कहुडे	kahadē	1	\rightarrow	कहड्डे	click here	kahadē	1

2. If it's a single-line (or multiple-line) bundle of ambiguities (left and right side), and one of them is incorrect, but the other side is not the correction for it, you can right-click on the select button in the center to split the one row into two:

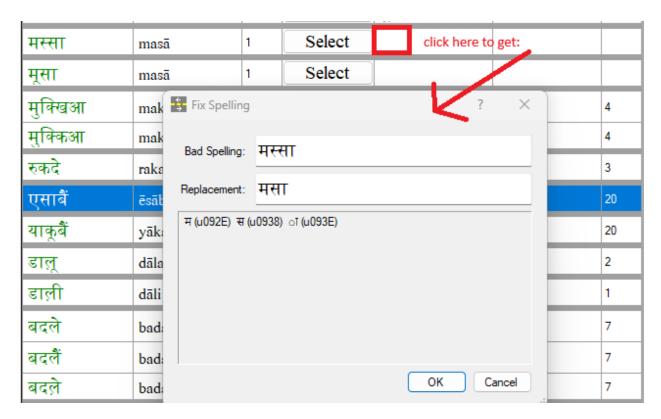
	•			S.	•	
मस्सा	masā	1	Select	मूसा	masā	1

fake example, but suppose that this left-hand word was misspelled (suppose it should have been a single 's') but the right-hand word is unrelated. The solution is to right-click on the center button (that says 'Select') and that will split the one line into two lines:

After right clicking on the 'Select' button, you get:

मस्सा	masā	1	Select		
मूसा	masā	1	Select		

• Finally, you can right-click on the right-hand side Vernacular column (whether it is empty as here or if there is a word in the cell) to bring up the correction dialog:



Additional Notes:

- As you make selections in each row whether one side is the correction of the
 other or both sides are valid words it is best to work through all the ambiguous
 bundles before closing the dialog. And you should save your work often (by
 clicking on *File*, *Save*). Saving the work will remove the processed rows from the
 table. If you close the dialog (even if you do *File*, *Save and Exit*), it will reload all of
 the words that you've already processed in case you want/need to change your
 answer.
- It is important as you process a word as being the correction of another word that the bad spelling word is always bad. When you eventually go to correct the spelling of documents (e.g. the SFM documents you loaded other otherwise), it will always change the bad spelling for to the good spelling form. So you don't want to use this to correct two distinct words that you realize are different words and need to be separated.
- Again, contact <u>silconverters support@sil.org</u> if you intend to use this tool, to get further guidance (I am not writing every suggestion for how you might use the results of this process, since I wrote this for my own needs and if no one else ever uses it, I'd rather not spend the time writing it up:-)