Fieldworks Language Explorer (FLEx) from SIL International

Reviewed by Lynnika Butler and Heather van Volkinburg
University of Arizona

1. OVERVIEW. Fieldworks Language Explorer (FLEx) is a new language documentation and analysis program that purports to combine the best features of Shoebox/Toolbox and LinguaLinks (all produced by SIL International). Like its predecessors, FLEx contains modules for documenting the lexicon and “texts” (any collection of utterances) of a language, as well as modules for describing the grammar, semantics, and other aspects. The lexicon can be used to create a dictionary, and text entries can be interlinearized to produce annotated morphological parses of utterances. Overall, our team has found the software to be more powerful, more visually appealing, and fairly easy to learn compared to Shoebox/Toolbox and LinguaLinks. However, FLEx has a number of minor to serious deficiencies that we feel need to be addressed in future versions.

Figure 1: Lexicon Edit view

2. NETWORKING AND COLLABORATION. Our team has created a sizeable digital database of the dormant Mutsun language from archival records, for use in revitalization as well as linguistic analysis. We recently switched from LinguaLinks (we previously used Shoebox and Excel) to FLEx for several reasons that may be relevant to other documentation
projects as well. First and most important for our team, FLEX can be networked to allow multiple collaborators to view and edit the database simultaneously. This had become an urgent problem in LinguaLinks, since at least three people needed to work on different aspects of the Mutsun database at the same time, and there was no foolproof way to merge everyone’s work into a single file. Networking in FLEX is fairly simple and has worked well for us thus far, although crashes sometimes occur when two people simultaneously attempt to edit the same object in the database. Fortunately, work is generally not lost in a crash.

3. PRICE, SUPPORT AND LONG-TERM DATA PRESERVATION. Several other advantages also influenced our decision to switch to FLEX. First, unlike LinguaLinks, the program is free, which makes it easily available to small documentation projects and language communities that may have little or no funding for software. Second, FLEX is being actively developed, while LinguaLinks is no longer supported, making it a liability in spite of some very sophisticated features. The FLEX software developers have been helpful and responsive when we needed assistance to convert our data from LinguaLinks, and as occasional problems or questions have arisen since. Finally, data in FLEX can optionally be backed up to XML, ensuring that many years of work do not end up in an incompatible legacy format when FLEX eventually becomes obsolete.

4. BULK EDITING. One of the much-touted features of FLEX is its Bulk Edit function, which allows one-step editing of all or part of the lexicon. This is an important advance over Shoebox/Toolbox and LinguaLinks, in which lexical entries had to be edited one at a time. In the lexicon window of FLEX, a user can select “Bulk Edit” for lexical entries, senses (roughly, glosses of an entry), or reversal entries (headwords in the gloss language). Some changes that we have made using this feature include: changing several incorrectly labeled morphemes from “prefix” and “suffix” to “proclitic” and “enclitic,” respectively; stripping off a particular suffix from a subset of verb stems; and recording consonant/vowel patterns of lexical items by bulk copying all lexemes into the “CV pattern” column, then replacing all vowels with “V” and all consonants with “C” (useful for syllable analysis, etc.). A minor frustration is that a few fields, such as “Examples” (sentences illustrating the use of a lexical item), can be displayed in the Bulk Edit view, but are not actually available for bulk editing.

5. CUSTOM FILTERING. The Bulk Edit, Lexicon Edit, and Lexicon Browse views all allow column-by-column filtering of the data, which is more intuitive than the complex filters that had to be created in LinguaLinks. As an example, in FLEX if we want to display only main forms that are nouns and that begin with the letter “c,” we can filter the Entry Type column for “main,” the Grammatical Info column for “noun,” and the Headword or Lexeme Form column for “c (at start).” In LinguaLinks, the same filtering required that we first create simple filters for “c”-initial items, nouns, and main entries, and then create a complex filter that combined all three simple filters.
6. DICTIONARY CREATION. While a dictionary can be created from the lexicon in FLEX, this is not nearly as straightforward a task as editing lexical data. We have found the “configure dictionary” tool to be less than intuitive to work with. However, the process was at least as difficult in LinguaLinks, so this is not a new problem; an hour or so of trial and error got our dictionary display into roughly the shape that we wanted.

One major flaw in FLEX exemplifies a complaint that our team has had about all linguistic documentation software since Shoebox: the lack of a full-featured gloss-to-vernacular dictionary output. We can use FLEX to create a rich Mutsun-English dictionary customized to display the information most useful to the Mutsun community, including part of speech, example sentences and translations, cross references, and semantic restrictions. However, the English-Mutsun (“reversal”) option provided by FLEX is simply an index of English headwords with their Mutsun translations, which language learners then have to look up in the Mutsun-English dictionary to fully understand.

Figure 3: Vernacular-to-gloss dictionary view
Because many documentation projects are also aimed at language revitalization, a fully customizable gloss-to-vernacular dictionary is at least as important a resource as the vernacular-to-gloss dictionary that can be produced by Shoebox, LinguaLinks, or FLEx, and we have long wondered why such a feature is overlooked in these programs. Our solution over the years has been to have a programmer write a script to reverse the relevant fields of the Mutsun-English dictionary to create an English-Mutsun version with the same display configuration as the original. However, not every documentation project has a programmer!

Another shortcoming of FLEx’s dictionary creation feature is particularly vexing in light of the program’s purported goal of enabling “professional quality research and publication, without a lot of expert consultation” (http://www.sil.org/computing/fieldworks/flex/overview.htm). FLEx is actually lacking a basic feature that both Shoebox/Toolbox and LinguaLinks offered: unlike the older programs, it cannot output dictionary data directly to a text file. The “Export Lexicon” function can be used to create an .xml or .db file of a configured dictionary; however, translating either of these into a printable (or web-viewable) dictionary requires at least an additional step, possibly many, and a certain degree of computer expertise.

To convert the XML to a text dictionary, a programmer would have to create a custom stylesheet to use with an XML converter (none appears to be offered by SIL). The .db file (called SFM for “standard format markup”), which looks like Shoebox, can be opened in Lexique Pro (also free from SIL) and then exported as .rtf, but this requires first telling Lexique Pro how to read the SFM field tags, setting up the language and writing system all over again, and a few other steps that are not self-explanatory. This is another big handicap for teams with limited computer expertise, and the FLEx help files are mysteriously silent on the subject of creating a print dictionary.

7. MANIPULATING TEXTS. In FLEx, texts can be viewed as “baseline” (vernacular text only) or “interlinear” (with configurable fields for morphological analysis, translations, and notes). We have found the text window somewhat slow to open, though this may be due in part to our large database (around 20,000 text entries). Navigating through texts is more difficult in FLEx than it was in LinguaLinks, because of the flat organization of texts.
(in LinguaLinks, sub-sections of texts could be organized hierarchically) and the lack of a single-entry text window. More alarmingly, our attempts to re-organize texts resulted in only the baseline texts being moved, while all interlinear information and translation/notes fields were lost. Finally, text entries can be modified only in the baseline view, but we often need to refer to our annotations in the interlinear view to determine what needs modifying. Toggling between the two views is slow and displays an entire section of text rather than the specific entry selected, requiring the user to hunt for the entry anew each time s/he changes views.

**Figure 5:** Text editor (interlinear view)

Another feature that we feel is lacking in FLEX’s text editor is the ability to search the translation and notes fields of text entries. In LinguaLinks, this could be accomplished indirectly via several steps; but in FLEX, the search tool in the text editor only finds strings that occur in the baseline text entries. As we have pointed out to the FLEX software development team, our project keeps copious notes and information pertaining to the original archival sources in the database, and we frequently have reasons to search for particular types of information in these fields. For example, we may wish to find all entries that the informant stated were another dialect of Mutsun, entries provided by a particular informant, or entries the original documenter or someone on our team marked as uncertain or requiring further analysis. Such a search tool would benefit any project which documents translations or other notes about their texts, whether elicited orally or copied from archives, and it is especially crucial for projects with more than a modest amount of text, for which manual searching is prohibitively time-consuming.

Because our team imported most of our text data already interlinearized from LinguaLinks, we have made minimal use of FLEX’s parser. However, the main observation thus far is that the parser has to be “loaded” before any string of text can be parsed. This
step has taken an excruciatingly long time (up to an hour) each time that we have attempted it. Again, this likely has to do with the large size of our database, but even with a text collection a quarter of the size of the Mutsun project’s, it seems very unwieldy. Parsing can be done semi-manually for individual strings of texts (a parse is chosen manually, but can be optionally applied to some or all identical strings elsewhere in the data); but one would prefer to attempt automatic parsing for large chunks of texts and leave the manual tool for correcting occasional mistakes.

8. CONCORDANCES. In the “Words” window of FLEX, a user can search for a string of text and display either all morphemes and text entries containing that string (“Concordance” view), or an analysis window for parsing the string (“Analysis” view). The main drawback to this feature is that it takes a long time to load (ten minutes or more for our project). Another flaw is that the search tool is designed to search from the left edge of a string, meaning that morpheme-level concordances are not possible. Our team has stumbled on a technique for searching for a string anywhere in a word (it also works in the Lexicon search tool), but this is not described anywhere in the FLEX help files: typing the “%” (the percent symbol) at the beginning of the string will find the string anywhere in a word. For example, searching simply for the string “pu” in our database will find only word-initial occurrences, ignoring all instances of the Mutsun reflexive suffix -pu, which is by definition non-initial. Typing “%pu” will find all occurrences of “pu,” including the reflexive suffix, but also all other strings containing the same letters, such as the verb stem cappu (“prick”).

One of the useful features of LinguaLinks was that once parses had been assigned to text entries, attested occurrences of a morpheme in the text could optionally be displayed in the lexical entry for that morpheme. Assuming correct parses, this meant that a morpheme-level concordance could be viewed in the lexicon. Also, these attested examples were hyperlinked to the full text entries in the text database, making it easy to navigate between the lexicon and text editors. The lack of such a feature in FLEX, especially given the sluggish startup of the “Words” window, makes this type of navigation much more time-consuming.

9. CONCLUSIONS. Our team has encountered a number of problems with the FLEX software that we hope will be addressed in future versions. However, it offers the ability to work collaboratively via a networked database, as well as the opportunity to use some powerful features such as bulk editing and concordance tools, making FLEX in many ways an improvement over Shoebox/Toolbox and LinguaLinks. Additionally, it is freely available, supported, and relatively easy to learn, making it potentially useful for language community members as well as linguists.
Pros: Networking capability (multiple collaborators can work simultaneously); Can be backed up as XML for long-term accessibility of data; Developers are helpful and willing to consider requests for improvements; Bulk edit features for lexicon; Easily customizable views; Work is not lost when crashes occur; Relatively easy to learn; Visually appealing; Free!!

Cons: Slow to open windows and process changes with a large database; Dictionary configuration is complex and not intuitive; Reversal index (gloss to vernacular language) is little more than a word list; No text export option for dictionary (XML or SFM only); No morpheme-level concordance (word-level only); Notes and translation fields cannot be searched; No direct linking between lexicon and texts as in LinguaLinks; No Macintosh version.

Primary functions: Quoted from http://www.sil.org/computing/fieldworks/flex/elicit and record lexical information; create dictionaries; interlinearize texts; study morphology

Platforms: PC

Open Source: Yes (however, our programmer gave up downloading the source code after it had run for 2 days!)
Available from: http://fieldworks.sil.org/License/SIL_open_source_license_frame.htm

Proprietary: Shareware

Available from: SIL International (http://www.sil.org/computing/fieldworks/flex/)

Cost: FREE

Reviewed version: 4.0.1

Application size: 256 MB

Documentation: http://fieldworks.sil.org/techsupportdocs.htm

Lynnika Butler
lgbutler@email.arizona.edu