Instructions for running CLAN MLU analysis
on Brown's Adam data

These instructions assume that you have downloaded the CLAN software and the Brown corpus from the CHILDES website, and decompressed them into a directory on your computer. The instructions are also written for the Mac version of the CLAN software. I hope and assume the PC version works the same way.

You'll probably need to change some of the program's default settings.

I needed to change the default LIB directory before clicking the HELP button did anything useful. On the Commands window, click the LIB button, then select the appropriate directory (in my case one folder deeper: CLAN: LIB), open this directory, then click Select Current Directory. You'll see the change noted next to the button on the Commands window.

You'll also need to change the WORKING directory to whatever directory you have Brown's data for Adam in. On my machine its in the directory brown corpus: adam. On the Commands window, click the WORKING button, then navigate to the directory in which you have the data, then click the Select Current Directory button.

In the Commands window, typing "dir" should now produce (in the CLAN Output window) a listing of all the adam data files:

```
> dir
adam01.cha   adam02.cha   adam03.cha   adam04.cha   adam05.cha   adam06.cha
adam07.cha   adam08.cha   adam09.cha   adam10.cha   adam11.cha   adam12.cha
adam13.cha   adam14.cha   adam15.cha   adam16.cha   adam17.cha   adam18.cha
adam19.cha   adam20.cha   adam21.cha   adam22.cha   adam23.cha   adam24.cha
adam25.cha   adam26.cha   adam27.cha   adam28.cha   adam29.cha   adam30.cha
adam31.cha   adam32.cha   adam33.cha   adam34.cha   adam35.cha   adam36.cha
adam37.cha   adam38.cha   adam39.cha   adam40.cha   adam41.cha   adam42.cha
adam43.cha   adam44.cha   adam45.cha   adam46.cha   adam47.cha   adam48.cha
adam49.cha   adam50.cha   adam51.cha   adam52.cha   adam53.cha   adam54.cha
adam55.cha

55 files, 0 directories

>
```

And at this point clicking HELP should produce this listing in the CLAN Output window:

```
=========================================================
>help

You can run any of the following CLAN programs:
chains        check        chip        columns      combo
compound      cooccur      dates       dist         dss
```

```
flo            freq          freqmerg       freqpos        gem
gemfreq        gemlist       keymap         kwal           lines
makedata       makemod       maxwd          mlt            mlu
modrep         mor           phonfreq       post           postlist
postmodrules   posttrain     rely           rtfin          saltin
statfreq       textin        timedur        vocd           wdlen
```

Or any of the following CLAN Plug-in programs:
SNDBoolets

Or any of the following CLAN utility programs:
```
chstring   combtier  delim     fixit     id         longtier  lowcase
repeat     retrace   uniq
```

Supplementary commands available:
```
accept     batch     cd        copy       del        dir        info
list       rename    rm        rmdir      ty         type
```

Command   Function

accept TEXT   allow TEXT type input files
accept ALL   allow any type of input file

batch FILENAME   runs a batchfile

cd ..   move up one directory
cd FolderName   change to a folder within current directory
cd HardDiskName: move to top level

copy [-luctqr] source destination
                  copy files
                  -l changes filenames to lowercase
                  -u changes filenames to uppercase
                  -cMSWD changes creator to MSWD, etc.
                  -tTEXT changes file type to TEXT
                  -q verification
                  -r recursive

del [-qr] filename   delete file
                  -q verification
                  -r recursive

dir   [-rl]   show contents of current directory
                  -r recursive
                  -l long format

info   list all current CLAN settings

list   lists the files currently under the @ sign

page filename   open up files for editing

ren [-luctqr] oldfile newfile
                  rename file
                  -l changes filenames to lowercase
                  -u changes filenames to uppercase
```

```
                    -cMCED changes creator to MCED for CLAN, -cMSWD for
Word,
etc.
                    -tTEXT changes file type to TEXT
                    -f rename even if the new file does already exist
                    -q verification
                    -r recursive

rmdir [-qr] dir        delete directory, if empty
                    -q verification
                    -r recursive

Left and right arrows move the cursor left and right on the command
line.
Up and down arrows rotate through the previous commands.


>
============================================================
```

Now we'll get ready to run a "mean length of utterance" analysis. Let's get the instructions first. Type MLU into the Command window:

The CLAN Output window will show the instructions for the MLU analysis:

```
============================================================

> mlu
MLU- Mean Length Utterance computes the number of utterances, morphemes
     and their ratio.
Usage: mlu [bS cS gF dN fS k pF rN re sS tS u yN zN ] filename(s)
+bS: make all S characters morpheme delimiters (default: -#~)
-bS: do not consider all S characters to be morpheme delimiters
+cS: look for unit marker S
+gF: exclude utterance consisting solely of specified words
+d : output in STATFREQ format. Must include speaker specifications
+d1: output data ONLY.
+fS: send output to file (program will derive filename)
-f : send output to the screen or pipe
+k : treat upper and lower case as different
+pF: define punctuation set according to file F
+rN: if N = 1 then "get(s)" goes to "gets", 2- "get(s)", 3- "get"
     4- recognize prosodic symbols in words, 5- no text replacement: [:
*]
     6- include repetitions: </> and <//>
+re: run program recursively on all sub-directories.
+sS: either word S or words in file @S to search for in a given input
file
-sS: either word S or words in file @S to be exclude from a given input
file
+tS: include tier code S
-tS: exclude tier code S
+u : merge all specified files together.
-u : compute result for each turn separately.
+y : work on TEXT format files one line at the time
+y1: work on TEXT format files one utterance at the time
+zN: compute statistics on a specified range of input data


>
```

```
=============================================================
```

Now we'll run this analysis on the first file of data from Adam. Doublclick the file adam01.cha. This will open a new Clan window, showing the transcript. This begins:

```
@Begin
@Participants:    CHI Adam Target_Child, MOT Mother, URS Ursula_Bellugi
            Investigator, RIC Richard_Cromer Investigator, COL
            Colin_Fraser Investigator
@ID:  en|brown|CHI|2;3.4|Male|normal|middle-class|Target_Child||
@ID:  en|brown|MOT|||||Mother||
@ID:  en|brown|URS|||||Investigator||
@ID:  en|brown|RIC|||||Investigator||
@ID:  en|brown|COL|||||Investigator||
=============================================================
```

Now you are ready to conduct an analysis. We won't set any parameters except the –f parameter, which sends output directly to the screen instead of to a file.

In the Command window, type:

mlu –f adamo1.cha

The CLAN Output window will report the results:
```
=============================================================

> mlu -f adam01.cha
mlu -f adam01.cha
Wed Jun 27 11:20:58 2001
mlu (13-Apr-2001) is conducting analyses on:
  ALL speaker tiers
****************************************
From file <Macintosh HD: Work More:CHILDES:brown
corpus:adam:adam01.cha>
MLU for Speaker: *RIC:
  MLU (xxx and yyy are EXCLUDED from the utterance and morpheme
counts):
      Number of: utterances = 5, morphemes = 17
      Ratio of morphemes over utterances = 3.400
      Standard deviation = 1.020

MLU for Speaker: *COL:
  MLU (xxx and yyy are EXCLUDED from the utterance and morpheme
counts):
      Number of: utterances = 0, morphemes = 0

MLU for Speaker: *URS:
  MLU (xxx and yyy are EXCLUDED from the utterance and morpheme
counts):
      Number of: utterances = 14, morphemes = 73
      Ratio of morphemes over utterances = 5.214
      Standard deviation = 2.932

MLU for Speaker: *MOT:
```

```
   MLU (xxx and yyy are EXCLUDED from the utterance and morpheme
counts):
      Number of: utterances = 447, morphemes = 1862
      Ratio of morphemes over utterances = 4.166
      Standard deviation = 2.274

MLU for Speaker: *CHI:
  MLU (xxx and yyy are EXCLUDED from the utterance and morpheme
counts):
      Number of: utterances = 1239, morphemes = 2600
      Ratio of morphemes over utterances = 2.098
      Standard deviation = 1.043
=========================================================
```

If you run the same analysis on the file adam55.cha, you'll find that MLU has increased from 1.043 (at age 2;3.4) to 2.555 (at age 4;10.0):

```
=========================================================
MLU for Speaker: *CHI:
  MLU (xxx and yyy are EXCLUDED from the utterance and morpheme
counts):
      Number of: utterances = 935, morphemes = 4240
      Ratio of morphemes over utterances = 4.535
      Standard deviation = 2.555
```