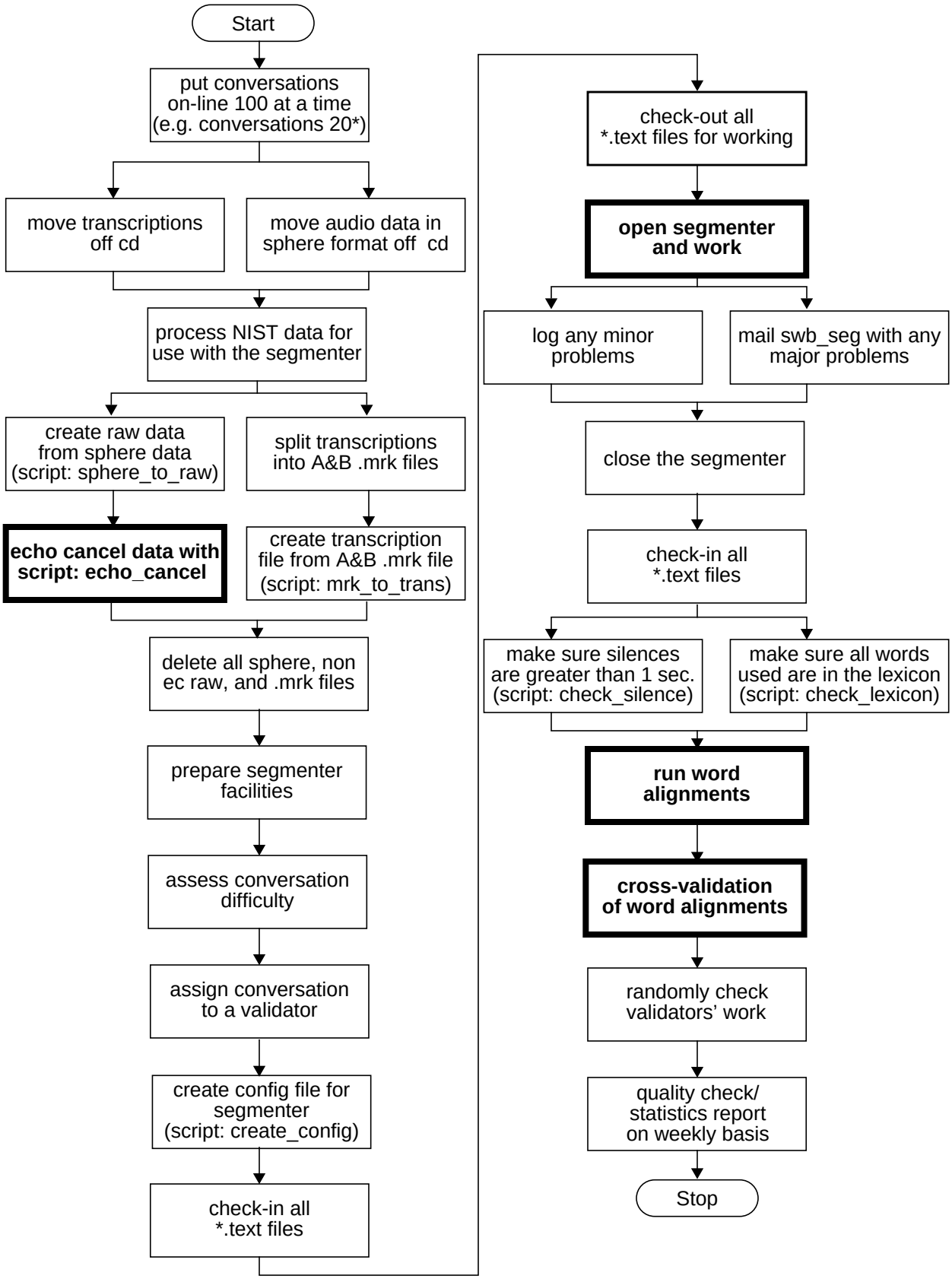
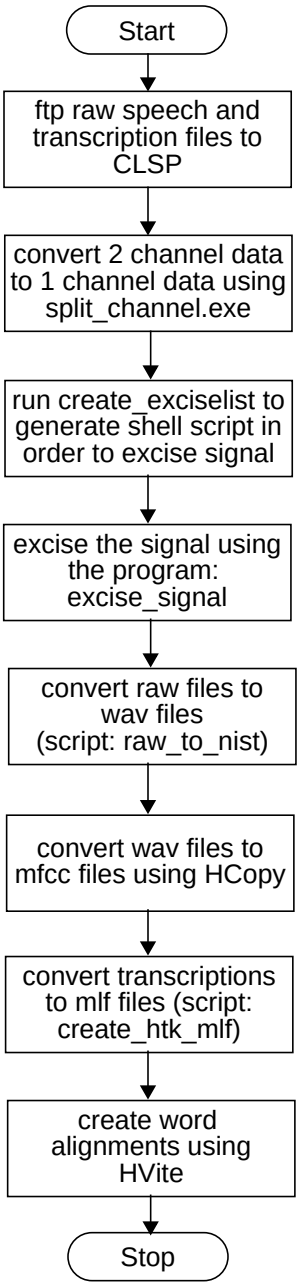

Start
put conversations on-line 100 at a time (e.g. conversations 20*)
move transcriptions off cd
move audio data in sphere format off cd
process NIST data for use with the segmenter
create raw data from sphere data (script: sphere_to_raw)
split transcriptions into A&B .mrk files
echo cancel data with script: echo_cancel
create transcription file from A&B .mrk file (script: mrk_to_trans)
delete all sphere, non ec raw, and .mrk files
prepare segmenter facilities
assess conversation difficulty
assign conversation to a validator
create config file for segmenter (script: create_config)
check-in all *.text files
check-out all *.text files for working
open segmenter and work
log any minor problems
mail swb_seg with any major problems
close the segmenter
check-in all *.text files
make sure silences are greater than 1 sec. (script: check_silence)
make sure all words used are in the lexicon (script: check_lexicon)
run word alignments
cross-validation of word alignments
randomly check validators' work
quality check/ statistics report on weekly basis
Stop

Work Flow Diagram: Word Alignments