**Linguistic variables**

**Pronunciation**

1. ð-stopping
   * Potential environments
     1. At least 50 tokens per speaker
     2. Maximum 3 tokens of any one lexical item (Carmichael)
     3. Don’t count if preceded by coronal stop (Labov 2001, Sph Int)
     4. Obviously not where standard has theta (“health,” “thank”)
     5. Accept any position (Carmichael)
        1. But need to check differently in spectrogram for initial vs medial vs final (Sph Int has deets for each) – attention to stop burst vs frication noise
   * Pronunciations: Fric/affric, stop, or neither
     1. Check audio by ear and spectrogram in PRAAT (Carmichael, Thomas SphOfCons); spectrogram examples Sph Intro pg 95
   * Process
     1. R script: from each audio transcript get list of unique TH words, export as csv
     2. In csv, remove theta words and repeats; separate into unique rows
     3. In audio extract text files, find 3 tokens of each viable word in appropriate contexts
     4. If more than three of a particular lexeme available, skip tokens in disfluent sections, when there’s other noise, etc. (may have to cut anyway)
     5. Whenever there’s enough tokens take the first from each of the three extracts; otherwise spread out as it works (e.g. if only in one extract, take near beginnning and end, not right next to each other, if possible)
     6. Check audio/spectrogram and indicate in csv
     7. 0 = deleted, 1 = fric/affric, 2 = stop

1. Verb-final “-ee”
   * Potential environments
     1. 50 main verbs per speaker
     2. Max 3 tokens of any one lexical item (Carmichael)
     3. Nothing that ends in a vowel
     4. One-syllable verbs only
     5. Bare present, after aux, or infinitive; not inflected for past/progressive/participle
   * Actual pronunciations
     1. Audio by ear and spectrogram in PRAAT
   * Process
     1. Still early stages – marking out appropriate verbs

**Grammar**

1. Pronominal apposition
   * Potential environments
     1. All finite declarative clauses (including 2nd-p subjects and existentials)
   * Actual uses
     1. Use of dislocated noun phrase and sentence with resumptive element coindexed to that noun phrase
     2. Aboutness requirement
   * Process
     1. Separated all extract audio files into clauses (rel clause, inf clause own line), following process used for BC/MM clause split as closely as possible
     2. Check each clause for dislocated noun phrase and coindexed resumptive element
2. Third-person self-reference
   * Potential environments
     1. First 50 clauses with self as subject
   * Actual uses
     1. Use of either “I” or own name or other reference to self in third-person (“I went” vs “Cudjo went” where “Cudjo” is the speaker)
     2. Only if speaking on own behalf, not including situations like repeating another person’s dialogue about the speaker
   * Process
     1. Use clause split from above
     2. Code each such clause for use of either “I” or own name/ref in third-person
3. Copula absence – is/are, am, was/were
   * Potential environments
     1. Clauses that *could* contain a copula (so not clauses that either cannot contain a copula or those that must contain a copula, see Rickford and Labov)
     2. Take all environments and separate out by person and tense
     3. If there’s enough, stop at first 150 potential environments
   * Actual use
     1. Whether the copula (am/are/is/was/were) is present in full, contracted, or absent
   * Process
     1. Use clause split from above
     2. Code for person: 1=first, 2=second, 3=third, 0=no subj (infinitive or fragment)
        1. Relative clauses counted for person of relativiser
     3. Code for tense: 1=past, 2=non-past (pres/fut/inf), 0=no verb (fragment, exclamation, etc.)
        1. Tense of verb rather than contextual timing of story (i.e. historical/storytelling present counted as present, even though story is in the past)
     4. Code for whether a sentence that can contain copula
     5. Remove sentences that must contain copula
     6. Code for full form, contraction, or absence
        1. 0 = absent, 1=contraction, 2=full form
4. “Ain’t” – present and past tense
   * Potential environments
     1. All clauses with negation
     2. Relatively uncommon variable, so take all uses from whole interview (not just extracts)
   * Actual use
     1. Use of “ain’t” or no such use
     2. Mark whether context suggests this is contraction of present-tense “has/is/are/am not” or of past-tense “was/did/had not”
   * Process
     1. Use clause split from above
     2. Code for clauses that contain a negation
     3. Code negation clauses for whether they contain “ain’t”
     4. Code “ain’t” uses for past vs present tense context
5. Infinitive “to” deletion
   * Potential environments
     1. Rather than first X environments, which would be difficult to define, better to get after every verb that shows such deletion in BC
     2. In BC: of 393 infinitives, 134 “to” deleted (34%)
        1. “Like,” “come,” “want,” “try,” “going,” “leave,” “go,” “get,” “need,” “start,” “send,” “tell,” “ought,” “pay,” “scared,” “run,” “hurt,” “stand,” “hurry,” “have,” “begin,” “stay,” “ask,” “suffer,” “gone,” and “pick up”
   * Actual use
     1. Whether the “to” is fully present, reduced, or absent
   * Process
     1. In csv, list of the BC verbs given above
     2. In interview extracts, mark each place where such a verb is followed by an infinitive
     3. In csv, tally of number of “verb + inf” constructions for each verb type
     4. Code for presence (full or reduced) vs absence
        1. Check by ear and also with spectrogram assistance in PRAAT for deleted vs very reduced, especially in contexts where reduction is common like “gonna,” “wanna,” “oughta”