Case Study: Citing archived resources in a Language publication

Paul Trilsbeek, Susan Kung, Mandana Seyfeddinipur
Publication in current form

- Publication by Anna Margetts in Language, December 2015: “Person shift at narrative peak”.
- Audio and ELAN files of Saliba-Logea examples available in The Language Archive, MPI Nijmegen
- From the acknowledgements:
  “Most of the data from the Logea dialect was collected by Carmen Dawuda between 2005 and 2008 as part of her Ph.D. research within the Saliba-Logea DoBeS project.”
•Current “citation”: footnote

“The database on which this study draws consists of texts collected in Papua New Guinea between 1995 and 2006. At the time the present study was conducted, the overall database included over sixteen hours of recorded speech by a wide range of speakers. The Saliba-Logea texts referred to in this article are available in the DoBeS archive to registered users. To register as an archive user go to http://corpus1.mpi.nl/ds/imdi_browser. The following texts include episodes of second-for-third-person markers: Boneyawa_12AH, Boneyawa_17AH, Boneyawa_18DL, Boneyawa_19DL, Boneyawa_20DL (1 and 2), Boneyawa_21DL, Boneyawa_23DZ, Conversation_01AN, Frogstory_02AZ, Kuhi_01AQ, Mouse7_02DA, Mouse7_06DX, Mouse0_05BQ, Taukulupokapoka_01AG, Tautolowaiya_01AG, and Wonu_01CN.”
Publication in current form

• Examples include Intonation Unit reference numbers which are also present in the ELAN files (also mentioned in a footnote).

• With some searching for file name and Intonation Unit, one can get to the examples in the archive fairly quickly
Example (from article)

(73) Ye-wane ‘Eh | tabu metai ku-lao-laoma …
     3SG-say INTRJ | PROH near.ADDR-LOC 2SG-RED-come
     ‘He said “Hey, don’t come toward me from there …’
     madai ye-kita-go.’
     lest 3SG-see-2SG.OBJ
     ‘otherwise it (the octopus) might see you.” ’

(Boneyawa_21DL_0029-31)

Search and browse in The Language Archive: [http://corpus1.mpi.nl](http://corpus1.mpi.nl)
Suggested Reference Citation

Suggested In-text Citation

• (Dawuda 2005a, 0029-31) instead of (Boneyawa_21DL_0029-31)
• Include range specification in in-text citation? In seconds, or reference IDs? (Dawuda 2005a, 00:00:55:846-00:00:58:591)?
• Option: include clickable link below each example?
Discussion Points

• Part identifier - intonational units, minutes & seconds, lines, etc.
• Citing researcher names instead of text names or codes (for analytics, altmetrics, etc.)
• When to cite the entire collection vs. parts of the collection. When citing an annotated media file, cite the “bundle”, or just the annotation file?
• Who is the “author”? Those who create the data should make that decision in the metadata.