

Background

The initial development of the *Carolinas Conversations Collection (CCC)* was supported by the National Library of Medicine from 2008-2011

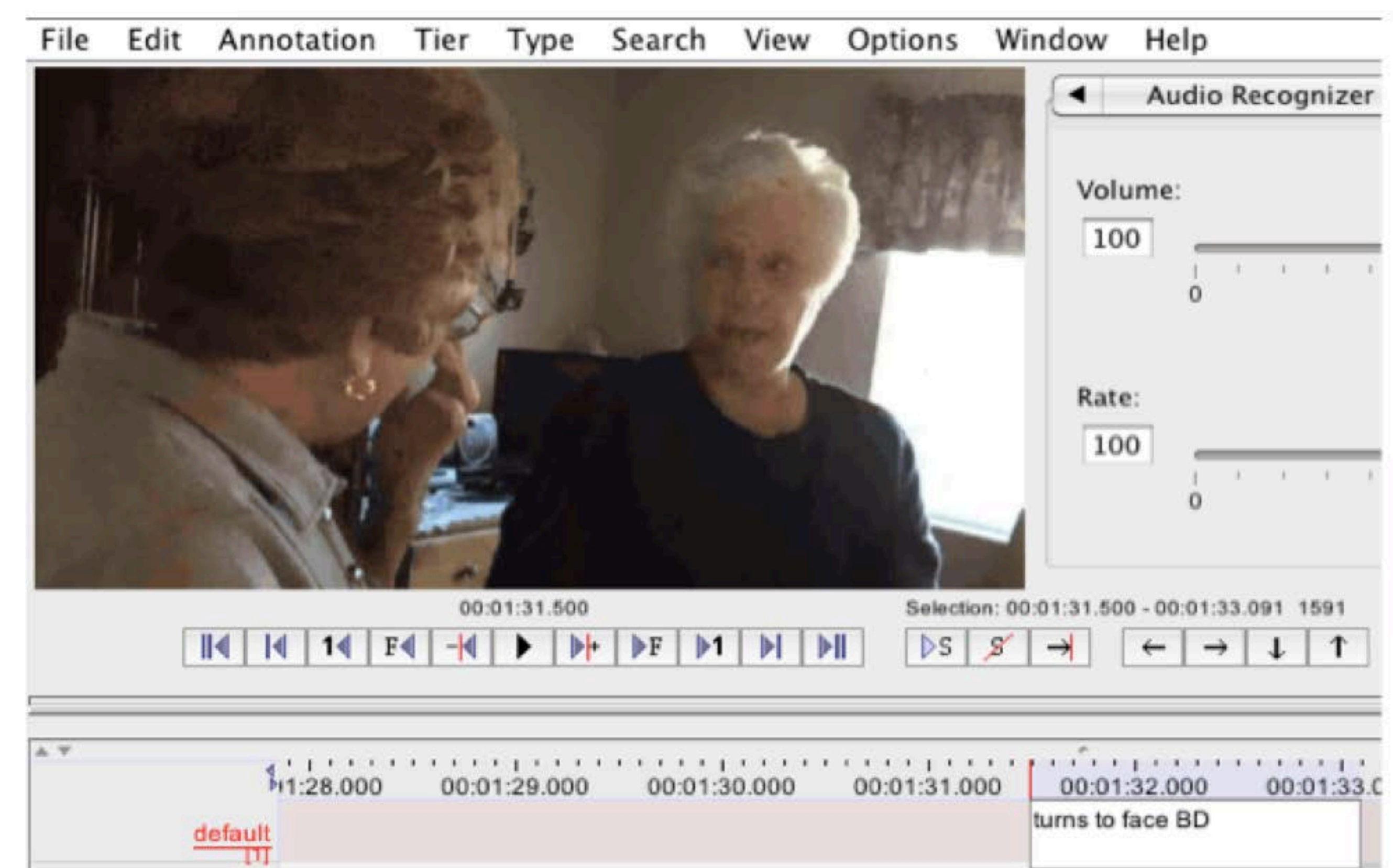


<http://carolinaconversations.musc.edu>

The core of this password-protected digital corpus presents audio, video and transcripts of men and women over 60 with and without dementia, talking about daily living, health and well being. The CCC portal allows access and retrieval of recordings in WAV or MP3, transcripts without visible markup, and some audio-synchronized transcripts which are web-browsable for online search, retrieval and analysis from word to sound signal.

name	transcripts	corpus	Role	Gender	Age range	Birthplace - City	First Language	Residence - State	Educational level	Condition or disease
Mr. Adrian	15	CCC	interviewee	M	70-79	en	South Carolina	13-16 years		Cancer/Prostate arthritis,
Ms Addison	1	CCC	interviewee	F	70-79	en	South Carolina	13-16 years		Diabetes HeartAttack Hy
Ms Ackland	11	CCC	interviewee	F	60-69	en	South Carolina	13-16 years		Diabetes Macular Degen
Ms Acton	5	CCC	interviewee	F	60-69	en	South Carolina	13-16 years		Hypertension
Ms Addison	14	CCC	interviewee	F	70-79	en	South Carolina	13-16 years		Diabetes
Ms Affleck	24	CCC	interviewee	F	70-79	en	South Carolina	13-16 years		Diabetes chronic liver fail
Ms Alan	22	CCC	interviewee	F	70-79	en	South Carolina	13-16 years		Hypertension shingles, de
Ms Appleby	13	CCC	interviewee	F	70-79	en	South Carolina	13-16 years		Diabetes HeartAttack Hy

Initially, the CCC held 80 transcripts of persons with dementia (PWD) with 33 involved in multiple conversational interviews. Donations made to this cohort are on-going: from 2008-13, an additional 315 conversations with PWD were collected and partially transcribed and another 120 have been donated since 2014. Currently, the central corpus has 875,739 words, and 692 transcripts with just over 800 hours of transcribed recordings, or 48,000 minutes. The CCC now also includes several associated corpora: Ratté's Mexican and Ecuadorean Spanish conversations with PWD, van Ravenstein's collection from low-income citizens aging in place and Wright's multi-faceted corpus of normally-aging persons, 30-90.



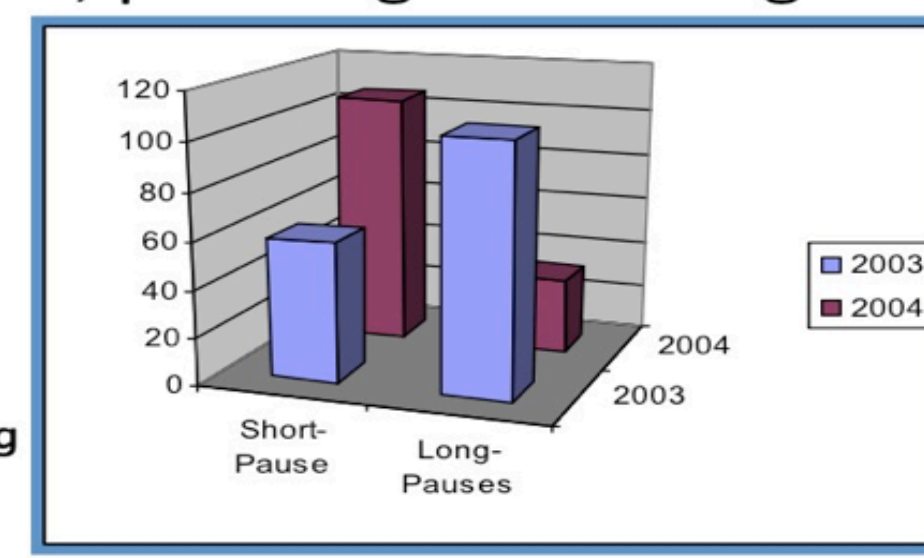
The CCC files can be imported into Elan and other software for multimodal analysis; the management package, LaBBCAT (<http://onzeminer.sourceforge.net/>) is linked to Praat, CELEX, and the Stanford parser for online analysis.

Pauses

Maintaining speaker identity and increasing social interaction both depend on language used by/to older speakers. Length and frequency of pausing plays a key part in how a person's fluency is perceived.

Pauses, planning and timing

In moving from early to moderate dementia, pauses shift function from word-finding to story-finding; timing gives way to planning at story-level



Pauses, Yr 2: planning for word, story, reformulation

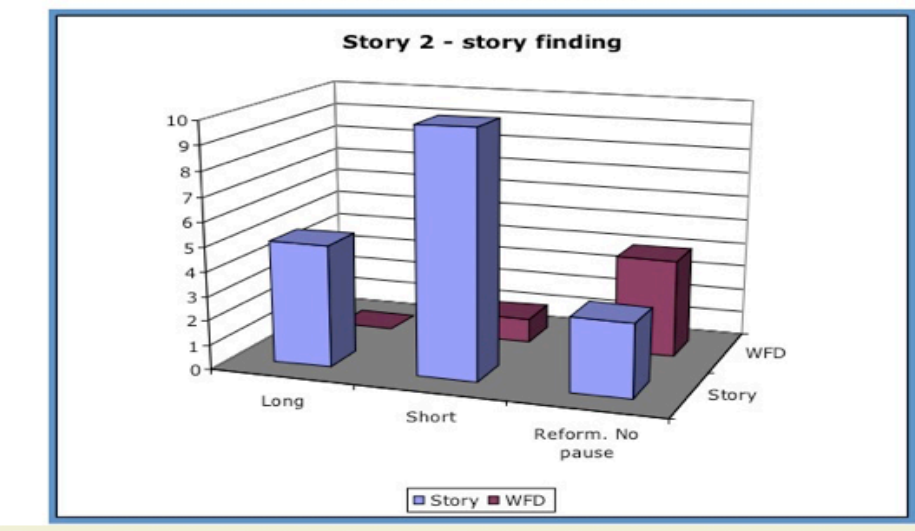
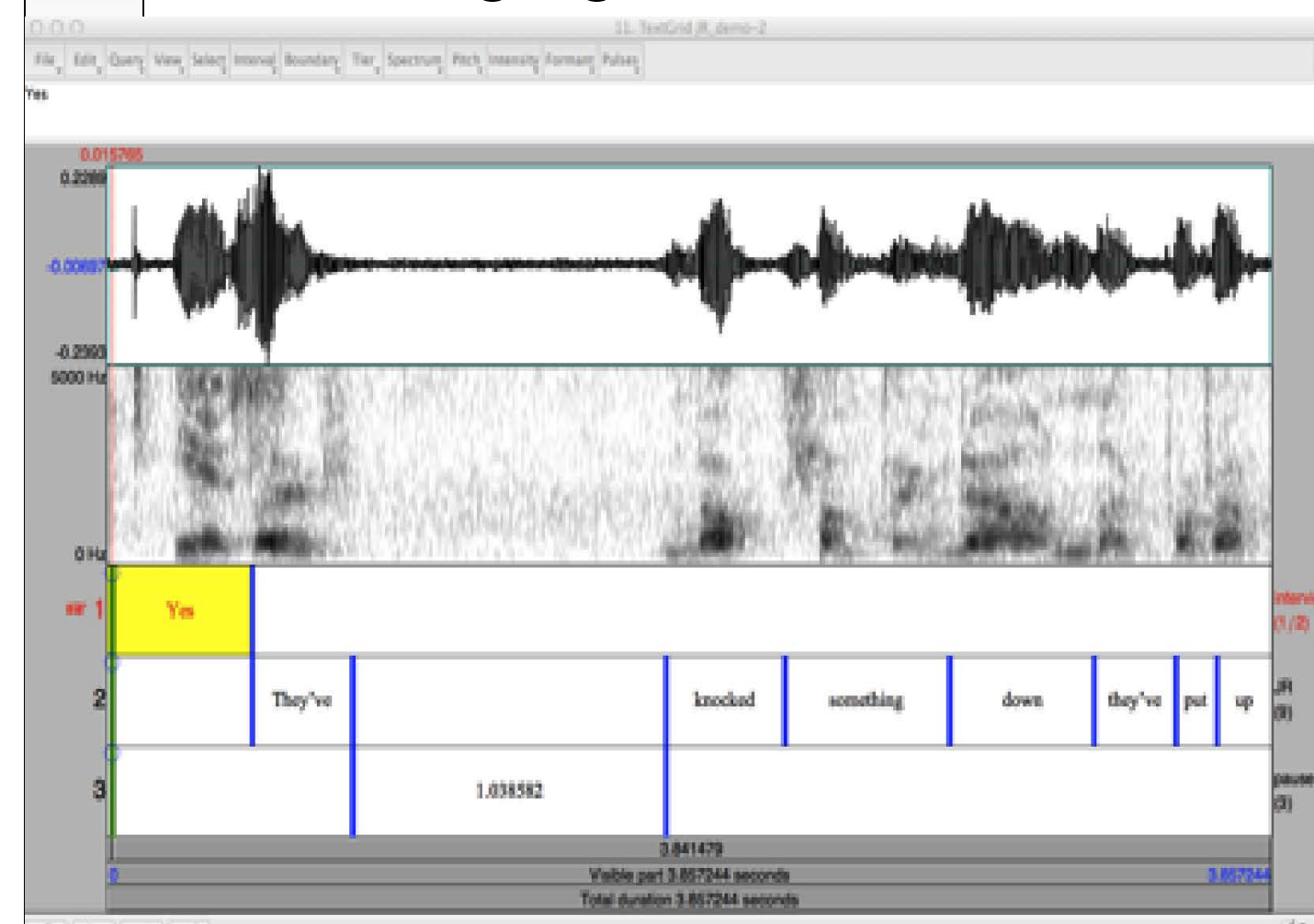


Figure 2: Davis & Maclagan 2009

Linguistic features such as pauses and prepositions contribute to our understanding of planning and of spatial and temporal concepts and usage, which are often linked to formulaic language used as pragmatic compensation for interaction maintenance. Such features require more systematic investigations of patterns to identify their relevance to language use in aging and classification of particular impairments.

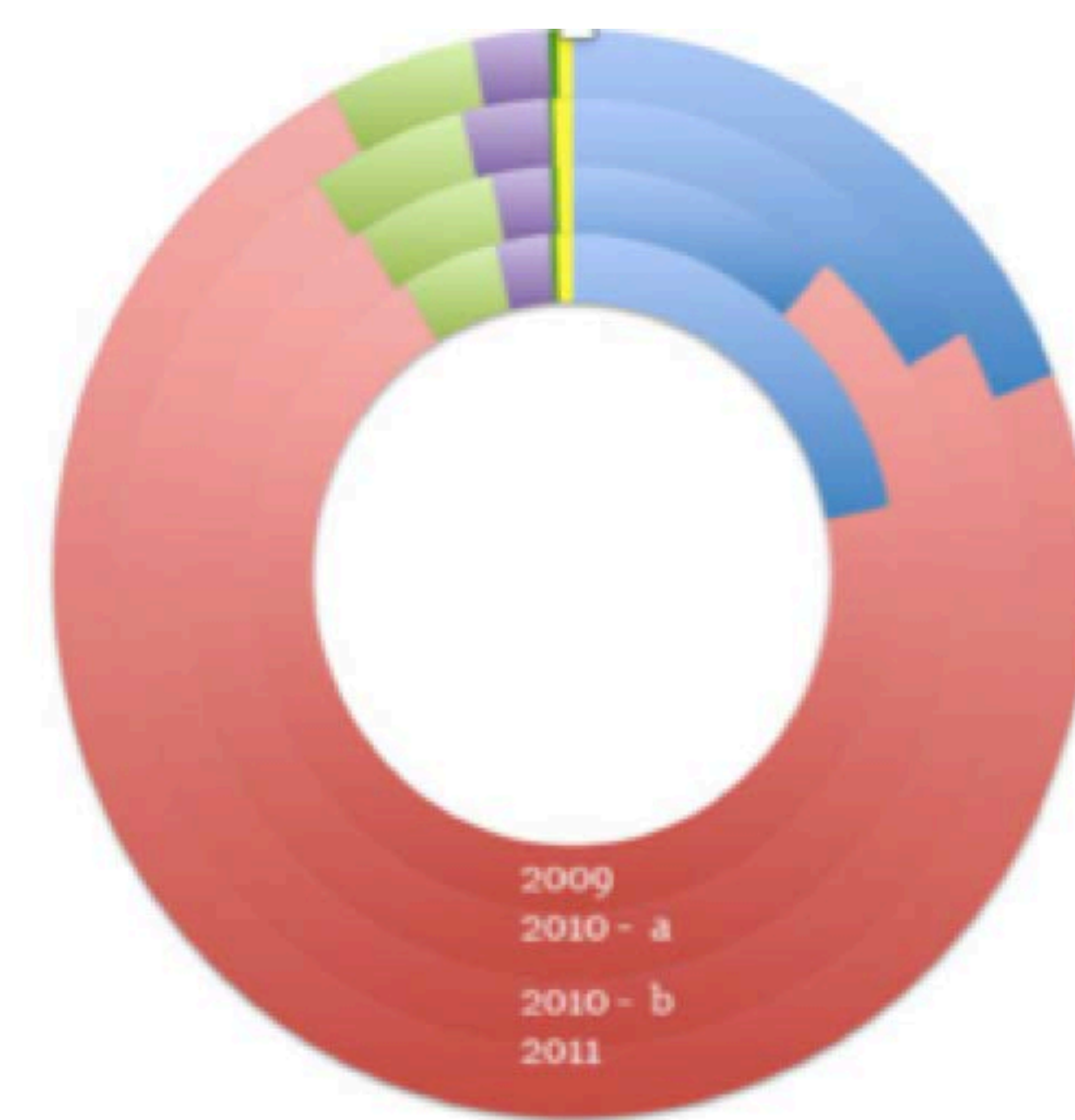


Screen shot, JR's pauses (Davis et al 2011).

Guinn and Habash (2014) sampled 80 CCC conversations between 31 PWD and 57 unimpaired partners. They found that POS tags and measures of lexical diversity were less useful than measuring filled pauses, repetitions and incomplete words. Pause length discriminated between persons with and without dementia in a walking-talking study focused on question types

Prepositions

The donut represents simple and complex prepositions in conversations across 3 successive years (n=11) with one woman as she aged from 85 to 88 and progressed from mild to early moderate dementia. It suggests an increase in compensatory strategies connected to a decrease in her ability either to retrieve story chunks, gist, or words. Our current investigation of reported speech in dementia conversation finds that it primarily occurs in scaffolded (Hyden & Orulv 2010) narratives as opposed to four other types of narrative format we identify as retained in early moderate dementia (Davis & Maclagan 2016). The collection invites your use: there is so much more to find.



Boersma, P. and Weenink, D. (2009) Praat: Doing phonetics by computer. [Computer program, version 5.106] <http://www.fon.hum.uva.nl/praat/> accessed 2 February, 2017.

Davis, B. and Maclagan, M. (2009) Examining pauses in Alzheimer's discourse. *American Journal of Alzheimer's Disease and other Dementias*, 24 (2): 141-154.

Davis, B., Maclagan, M., Karakostas, T., Hsiang, S. and Shenk, D. (2011) Watching what you say: walking and conversing in dementia. *Topics in Geriatric Rehabilitation*, 27 (4): 268-277.

Davis, B. and Maclagan, M. (2017, submitted) Narrative and aging: Exploring the range of narrative types in dementia conversation. *European Journal of English Studies*.

Guinn, C., Singer, B., Habash A. (2013) A comparison of syntax, semantics, and pragmatics in spoken language among residents with Alzheimer's disease, 2014. Proceedings of IEEE Symposium Series on Computational Intelligence <http://ieeexplore.ieee.org/document/7007840/>

Hyden, L. & Orulv, L. (2009). Narrative and identity in Alzheimer's disease: A case study. *J Aging Studies*, 23: 205-214.