A Swedish Cookie-Theft Corpus

Dimitrios Kokkinakis[1], Kristina Lundholm Fors[1], Kathleen Fraser[1], Arto Nordlund[2†]
Department of Swedish[1]; Department of Psychiatry and Neurochemistry[2]
University of Gothenburg, Sweden
{dimitrios.kokkinakis; kristina.lundholm; kathleen.fraser}@svenska.gu.se
{arto.nordlund}@neuro.gu.se

Abstract

Language disturbances can be a diagnostic marker for neurodegenerative diseases, such as Alzheimer's disease, at earlier stages. Connected speech analysis provides a non-invasive and easy-to-assess measure for determining aspects of the severity of language impairment. In this paper we focus on the development of a new corpus consisting of audio recordings of picture descriptions (including transcriptions) of the Cookie-theft, produced by Swedish speakers. The speech elicitation procedure provides an established method of obtaining highly constrained samples of connected speech that can allow us to study the intricate interactions between various linguistic levels and cognition. We chose the Cookie-theft picture since it's a standardized test that has been used in various studies in the past, and therefore comparisons can be made based on previous research. This type of picture description task *might* be useful for detecting subtle language deficits in patients with subjective and mild cognitive impairment. The resulting corpus is a new, rich and multi-faceted resource for the investigation of linguistic characteristics of connected speech and a unique dataset that provides a rich resource for (future) research and experimentation in many areas, and of language impairment in particular. The information in the corpus can also be combined and correlated with other collected data about the speakers, such as neuropsychological tests, brain physiology and cerebrospinal fluid markers as well as imaging.

Keywords: Speech elicitation corpus, Cookie-theft stimuli, dementia, linguistic features, language impairment

1. Introduction and Background

Elicitation of spontaneous speech and narrative discourse samples from (young and/or adult) individuals both healthy as well as with a variety of mental and cognitive impairments, at various stages of the pathology and with a variety of stimuli, have been developed and studied for several decades (e.g., Bottenberg et al., 1987; Bryant et al., 2016). Such systematic form of language production can be a valuable procedure for describing and even quantifying the degree and severity of e.g. cognitive decline, which most often[1] incorporate a single picture description, a picture sequence description (e.g., the *Dog Story*; de Lira et al., 2011) or a personal opinion on an event or idea. Even work on normative data has been reported for use in various narrative discourse tasks. For instance, based on a large control sample, Richardson & Dalton (2015) created main concept checklists (information content units) for three semi-spontaneous discourse tasks, a picture sequence narrative *Broken Window*, a storytelling *Cinderella*, and a procedure *Peanut Butter and Jelly*. For a similar purpose, Catricalà et al. (2017) created their own original picture *Summer Time* used as normative data on 134 Italian subjects pooled across homogeneous subgroups for age, sex, and education.

Nonetheless, the single black and white picture of the *Cookie-theft* from the Boston Diagnostic Aphasia Examination Battery (BDAE; Goodglass et al., 2001) is one of *the* most used stimuli for the assessment of language production. Other images, such as the *picnic* picture (Weissenbacher et al., 2016) from the Western Aphasia Battery (Kertesz, 1982) or the tales of *Cinderella, Snow White, the Seven Dwarfs* and the *Little Red Riding Hood* (Machado Lima et al., 2014; Silveira & Mansur, 2015) have also been used in various studies but to a much lesser extent; for an extended review see Bryant et al. (2016) and Boschi et al. (2017).

In this paper we describe a Swedish Cookie-theft corpus developed in the "Linguistic and extra-linguistic parameters for early detection of cognitive impairment" project. The project aims to adapt, develop and test methods that in isolation have shown promising outcomes on tasks related to early detection of dementia, differentiating between various dementia types and controls and explore the cognitive processes that underlie certain forms of spoken language production.

2. The Cookie-Theft Stimulus

The single structured stimuli of the Cookie-theft picture depicts a familiar household event scene in a kitchen which includes a woman drying some dishes, unconcerned by the overflowing of water from the sink, and the notable feature of a child (boy) standing on a stool about to tip over, stealing cookies (thus the name of the picture) off a high shelf. The stimuli is considered an ecologically valid approximation to spontaneous discourse (Gilles et al., 1996) and have been widely used to elicitate speech from speakers with different types of language and communication disorders, including numerous discourse tasks for the study of various types of dementia. In contrast to spontaneous speech the Cookie-theft allows control for content and context. Moreover, the figure is often used

† Deceased November 2017.

[1] Other spoken language elicitation procedures include the Wechsler's Logical Memory (Wechsler, 1997), an immediate and delayed story recall test.

because of the clear, straightforward drawings that reduce ambiguity and lessen the effect of memory problems. The picture allows the variation between speech-styles to be reduced, and also "minimizes confounds in analysis due to the controlled nature of the speech content" (Williams et al., 2010).

2.1 Previous Work Based on the Cookie-Theft

Narratives based on the Cookie-theft picture have been a source of knowledge for clinical and experimental research worldwide which also enables potential cross-linguistic comparisons. The Cookie-theft stimuli has been used to elicit written narratives, and this can be especially useful when evaluating people with Alzheimer's, as written language has been shown to be impaired even at the early stages of the disease. E.g., Forbes et al. (2004) reported significantly worse written discourse production at the mild and moderate stages of AD compared to healthy controls; while Kavé & Goral (2016) confirmed that individuals with Alzheimer's produce a lower proportion of nouns out of all other words as well as a higher proportion of pronouns out of all other words. Although various datasets based on the Cookie-theft picture are written ones (Croisile et al., 1996 for French and Pekkala et al., 2013 for English); other combinations exist, such as both written and verbal (Groves-Wright et al., 2004); as well as audio recordings, with or without transcriptions which have been collected and described for several languages (e.g., English, Ripich et al., 2000; Swedish, Tyche, 2001; German, Bschor et al. (2001); Spanish, Cuetos et al., 2007; Japanese, Choi, 2009; Norwegian, Lind et al., 2009 and Hebrew, Kavé & Dassa, 2017). Moreover, the Cookie-theft picture has been analysed at various linguistic levels (Kavé & Levy, 2003; de Lira et al., 2014; Yancheva & Rudzicz, 2016) and for different purposes (e.g., aphasia examination, Gilles et al., 1996; Ash et al., 2013; acquired brain injuries, Hux et al., 2008; and normative data collection, Forbes-McKay & Venneri, 2005).

2.2 Cookie-Theft and Information Content Units

Information content units, ICUs, refers generally to various predefined image concepts/objects/facts and actions describing a picture. With respect to the Cookie-theft picture content there exist several inventories that try to capture all possible semantic or information content units, containing between 7-25 ICUs. Mackenzie et al. (2007) uses seven such ICUs ('woman doing dishes', 'sink overflowing', 'boy on stool', 'children stealing cookies', 'girl reaching for cookie', 'stool falling', 'woman not noticing') while Kavé & Levy (2013) proposed 25[2], divided into four semantic groups; namely: *subjects* ('boy', 'girl', 'woman'), *places* ('kitchen', 'exterior seen through the window'), *objects* ('cabinet', 'cookies', 'counter', 'curtain', 'dishes on the counter', 'faucet', 'floor', 'jar', 'plate', 'sink', 'stool', 'water', 'window') and *actions* ('boy taking the cookie', 'boy or stool falling', 'woman drying or washing dishes/plate', 'water overflowing or spilling', 'the girl asking for a cookie', 'woman unconcerned by the overflowing', 'woman indifferent to the children'). If used, synonymous expressions were accepted for the same concept (e.g. 'mammy' and 'mother' and also 'woman'); even the wrong word referring to a piece of information is accepted (e.g., 'ladder' instead of 'stool'; 'shelf' instead of 'cabinet' or 'cupboard')

Moreover, these information units can be identified both with a top down approach (Croisile et al., 1996) or a bottom up one, e.g. using clustering word embeddings of the naming expressions (Yancheva & Rudzicz, 2016). Usually, subjects are given credit for mentioning the presence of a given ICU and thus this set of ICUs enables easy comparison across subjects and easy scoring of the predefined contents of the picture (Boschi et al., 2017). Sirts et al. (2017) applies the notion of semantic idea density[3] (SID) on the same image, in which each SID is a count of the pre-defined information content units and is computed by counting the number of ICUs mentioned in the text and then normalising them by the total number of word tokens.

2.3 Cookie-Theft Datasets

The most extensive dataset of the picture description publicly available is the DementiaBank[4] corpus, a part of the TalkBank project, collected between 1983-88; (Becker et al., 1994; MacWhinney et al., 2011). DementiaBank is a clinical dataset which consists of interview recordings and transcripts of English-speaking people with AD and a few with mild cognitive impairment (MCI), describing the Cookie-theft picture. This dataset has been used in various studies in the past, e.g. Fraser et al. (2016) achieved classification accuracies of over 81% in distinguishing between participants with AD and healthy controls by computing a number of linguistic variables from the transcripts, and acoustic variables from the associated audio files; Orimaye et al. (2017) used DementiaBank to learn syntactic, lexical, and n-gram linguistic biomarkers to distinguish the group of probable AD from the healthy group. Their best diagnostic model significantly distinguished the two groups using Support Vector Machines; while Masrani et al. (2017) used the DementiaBank data for domain adaptation and discourse features in an attempt to improve classification accuracy The Cambridge Cookie-theft corpus (Williams et al., 2010)

[2] The same 4 key categories, but with less ICU content (23), were identified by Croisile et al. (2016); 'dishcloth' was an additional object used by Croisile et al., while 'floor', 'counter' and 'faucet' were three objects used by Kavé & Levy (2013) but not used by Croisile et al.

[3] A complementary notion to SID is propositional idea density or PID (i.e. the number of any propositions or ideas expressed in the text, that is, the amount of information that is conveyed relative to the number of words used to encode it; Chand et al, 2010) which according to Sirts et al. (2017), is more suitable on datasets of spontaneous speech on free topics rather than standardised picture descriptions.

[4] DementiaBank: <dementia.talkbank.org/> Visited 20180102.

is another corpus consisting of samples by healthy individuals without brain-damage and patients recorded between the mid-1990s and 2009. The OPTIMA[5] (Oxford Project to Investigate Memory and Aging) transcripts of connected speech has been also used in several studies. OPTIMA is a longitudinal study of aging and dementia in a cohort of elderly, community living volunteers collected between 1998-2008. For instance, Rentoumi et al. (2014) used the OPTIMA Cookie-theft descriptions in machine learning text classification experiments, in order to assign the samples to one of two pathological groups (mixed vascular and AD and pure AD) on the basis of lexical and syntactic features as well as statistical and information theory characteristics, with high classification accuracy (75%). The Framingham Heart Study[6] is a longitudinal population-based study established in 1948, a number of dementia-free sub-cohorts have been followed since the mid-1970s for development of incident dementia (Seshadri et al., 2006). A neuropsychological battery administered three times to these participants included a written description of the Cookie-theft picture. The purpose of the Framingham Heart Study was to find an early language marker of AD by comparing lexical retrieval on these tests in participants with clear cases of AD compared to that of participants who did not. Finally, the WRAP (Wisconsin Registry for Alzheimer's Prevention) study sample (Johnson et al., 2017), with recordings and transcriptions of 264 participants (200 cognitively healthy and 64 with early MCI), was used for determining if participants with very early, subclinical memory declines were also showing declines in connected language. Based on these data, Mueller et al. (2017) showed that participants with early MCI status declined faster in features of speech fluency and semantic content than those who were cognitively stable.

2.4 The Swedish Cookie-Theft Corpus

The Swedish Cookie-theft corpus[7] described in this paper consists of data from three subject groups: those with subjective cognitive impairment (SCI), those with mild cognitive impairment (MCI), and healthy controls (HC). All participants were recruited from a longitudinal study in the region of West Sweden (see Section 3). Note that SCI, MCI, and AD are on a spectrum of disease progression. SCI is a common diagnosis in elderly people, sometimes suggested to be associated with e.g. depression or anxiety, but also a risk factor for dementia (Jessen et al., 2010). MCI is a well-defined prodromal state of dementia and appears to represent a transition between normal aging and early dementia (Ritchie & Touchon, 2010), in which a person has minor problems with cognition (e.g., problems with memory or thinking) but these are not severe enough to warrant a diagnosis of dementia or interfere significantly with daily life. Still, persons with MCI have difficulties which are worse than would normally be expected for a healthy person of their age. In clinical settings patients with MCI show rates of conversion to dementia of about 12% per year (Petersen et al., 2001).

3. The Gothenburg MCI-study and Related Ethical Issues

The Gothenburg mild cognitive impairment study (Nordlund et al., 2005; Wallin et al., 2016) conducts longitudinal in-depth phenotyping of patients with a wide range of cognitive impairment (i.e., from very mild to manifest dementia including, but also including cognitively normal controls) using neuropsychological, neuroimaging, and neurochemical tools. The study is clinically based and aims at identifying neurodegenerative, vascular and stress related disorders prior to the development of dementia. All patients in the study undergo baseline investigations, such as neurological examination, psychiatric evaluation, cognitive screening, magnetic resonance imaging of the brain and cerebrospinal fluid collection. At biannual follow-ups, most of these investigations are repeated. The Gothenburg MCI-study is approved by the local ethical committee review board (ref. L091–99, 1999; T479-11, 2011); while the currently described study is approved by the local ethical committee decision 206-16, 2016. The project aims at gathering a rather homogeneous group of participants with respect to age and education level. Written informed consent is obtained from all participants in the study while the exclusion and inclusion criteria are specified according to a predefined protocol.

4. Population and Socio-Demographic Characteristics

All subjects in the three different groups (HC, SCI and MCI), were native speakers of Swedish with adequate hearing and vision, no documented history of neurological or psychiatric illness, learning disabilities or e.g. alcoholism according to the inclusion and exclusion criteria pre-specified, aged 53-79 years old, while 53 of them (58%) were females. Demographic details and MMSE scores for all participants are provided in Table 1.

MMSE, the Mini Mental State Examination, is a paper-based questionnaire and a screening test used to evaluate cognitive function, where 30 points is the maximum, and 24 or above indicates normal cognition (Folstein et al., 1975). The difference in age between the groups is not statistically significant, as determined by a one-way ANOVA ($F(2, 88) = 2.082$, $p = .131$). When comparing number of years of education, a one-way ANOVA shows

[5] OPTIMA: <www.ndcn.ox.ac.uk/research/centre-prevention-stroke-dementia/resources/optima-oxford-project-to-investigate-memory-and-ageing> Visited 20180102.

[6] Framingham Heart Study: <www.framinghamheartstudy.org> Visited 20180102.

[7] We are aware of two more Swedish Cookie-theft collections. Tyche (2001) collected oral descriptions from 12 MCI and 12 healthy controls; while Cromnow & Landberg (2009) collected written samples from 96 healthy, Swedish controls, having Swedish as a mother tongue, absence of clinical manifestations of linguistic impairment and absence of neurological disease. This constitutes normative data from speakers without cognitive impairments and which its purpose was to establish norms.

that there is a significant difference between the groups ($F(2, 88) = 6.085$, $p = 0.003$). A post-hoc LSD test demonstrates that there is no significant difference between the HC and the MCI group, but that there is a significant difference between the HC and the SCI group ($p = 0.001$) and the MCI and the SCI group ($p = 0.02$). The MMSE score differs significantly between the three groups, as determined by a one-way ANOVA ($F(2, 88) = 16.978$, $p < 0.0001$). Post-hoc LSD test show that subjects in the MCI group have significantly lower MMSE scores than the HC group ($p < 0.0001$) and the SCI group ($p < 0.0001$), whereas there is no difference between the SCI group and the HC group. To summarise, there is no difference in age between the groups, but the subjects in the SCI group have a longer education than the MCI and the HC group. Finally, the subjects in the MCI group score lower on the MMSE than the subjects in the HC group and the SCI group.

	HC (n=36)	SCI (n=23)	MCI (n=32)
Mean Age_{years}	67.9 (sd 7.2)	66.3 (sd 6.9)	69.9 (sd 5.7)
Mean Edu_{years}	13.2 (sd 3.4)	16.1 (sd 2.1)	14.1 (sd 3.5)
Female	23	14	16
Male	13	9	16
MMSE	29.6 (sd 0.61)	29.5 (sd 0.90)	28.2 (sd 1.41)

Table 1: Demographic information for the three groups of participants. The MiniMental State Exam (MMSE) is a test of cognitive status and has a maximum score of 30.

The population was selected from an ongoing epidemiological study in the region, see Section 3. At the start of the study, each participant was given information about the purpose of the evaluation, and written informed consent was obtained in accordance with the protocol approved by the local ethics committee. Note that the mean result on the MMSE is above 24 for all groups (a cutoff of ≤ 24 has been proposed for cognitive impairment, with a score between 25-27 indicating possible cognitive impairment which should be further evaluated) which means that all test within the range of normal cognition; both the MCI and SCI diagnosis are indicative of subtle cognitive impairment and it is expected that the subjects perform relatively well on tests such as the MMSE.

5. Data Collection

During the presentation of the Cookie-theft stimuli the subjects are asked to describe everything that can be observed in the picture, and the narrative is recorded. For the task, the original label on the image's cookie jar is translated and substituted from the English "COOKIE JAR" to the Swedish label "KAKBURK". The samples are recorded in an isolated environment and the whole task is designed to evoke a monologue by the participant. The instruction given to the subject was: "Tell me everything you see going on in this picture, describe objects and events. You can go on as long as you prefer and you will be not interrupted until you indicate that you do not have anything more to say". Descriptions were recorded and the recordings were terminated when the subject indicated verbally or non-verbally (by becoming silent for a longer period of time) that the narrative was completed.

	HC (n=36)	SCI (n=23)	MCI (n=31)
Shortest description	29s	41s	21s
Longest description	274s	635s	361s
Mean duration	92.3s	123s*	102.3s
Total duration	3324s	2829s	3173s

Table 2: Duration of the Cookie-theft samples (*the mean duration for the SCIs is 99.7 seconds if we exclude one extreme outlier description of 635 seconds)

For the audio capture of the narratives we used a H2n Handy recorder and the resulting audio files were saved and stored as uncompressed audio in .wav 44.1 kHz with 16-bit resolution. The recording is carried out in an isolated environment at the University lab in order to avoid noise. A speech pathologist and computational linguist were present during the recording sessions, providing all subjects with identical instructions according to a pre-defined protocol. Information about the length of the spontaneous speech samples is provided in Table 2.

6. Transcriptions and Linguistic Analysis

Based on the digitized speech waveform of the Cookie-theft audio recordings, manual transcriptions are produced, that are semi-automatically aligned with the audio signal. During speech transcription, special attention is also paid to non-speech acoustic events including speech disfluencies consisting of filled pauses a.k.a. hesitations (such as "um" and "uh"), false-starts, word fragments, repetitions as well as other features, particularly non-verbal vocalizations such as laughing, sniffing and coughing. A basic transcription manual, with the various conventions to be used, is produced which helped the human transcribers accomplish a homogeneous transcription. For instance, all numerals should be written out as complete words, while symbols, such as XML-tags, are used for the encoding of pauses or transcriber's comments. Furthermore, for the transcription the Praat application (Boersma & Weenink, 2013) is utilized; using a 2-tier text grid configuration, one for broad orthographic transcription (standardized spelling with manual utterance segmentation) and one for narrow orthographic transcription, with maintained spoken language phenomena, such as partial words, see Figure 1.

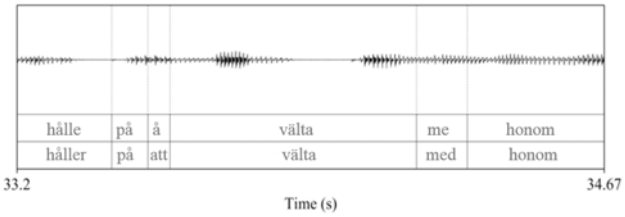

Figure 1: Example of transcription of the collected data.

7. Linguistic Processing, an Overview

In total, there are 19285 tokens in the sample (HC: 8770; MCI: 5724; SCI: 4791) and 1573 word types (HC: 1077; MCI: 750; SCI: 734). The transcriptions have been processed with basic Natural Language Processing (NLP) tools for Swedish which includes tools for lexical analysis (lemmatization and multiword expression identification), part-of-speech annotation, syntactic parsing (both dependency and constituent-based) and identification of information content units. Computational models that can be used to differentiate between the groups are typically based on various lexical, syntactic, semantic, and discourse characteristics. Here we provide a very general overview of the various features we currently extract from the corpus, both the audio signal (i.e. acoustic and prosodic) and the transcriptions (i.e. features that follow the lexicon-syntax-semantics continuum). Results of our linguistic analysis are reported elsewhere (e.g. Lundholm Fors et al., 2018; Fraser et al., 2018).

From the transcriptions we extract: (i) lexical distribution measures (such as type-token ratio, mean word length, long word counts but also the distribution of lexical and non-lexical fillers or disfluency markers such as "um", "uh", "eh") and out-of-vocabulary rate (Pakhomov et al., 2010); (ii) syntactic complexity markers (such as [context free] production rules; dependency distance; noun phrase average length and noun phrase density; (iii) psycholinguistic measures (such as familiarity); (iv) imformation units (such as content density). Measures that can give an indication of semantic deficits are important and likely to be present even in the early stages of MCI. A large number of acoustic features (such as speech rate, pause frequency, filled pauses, total pause duration; pitch and formants) have been proposed as relevant in the literature (Roark et al., 2011; Ahmed et al., 2013; Yancheva et al., 2015) and it pinpoints the importance of distinguishing between vocal changes that occur with normal aging and those that are associated with MCI (and SCI).

8. Conclusions and Future Work

This paper presented a new Swedish corpus of semi-spontaneous oral text production (connected speech) of the Cookie-theft picture by speakers with subjective and mild cognitive impairment and healthy age-matched controls. The main purpose of the corpus described in this paper is to use it for the identification of subtle language deficits in spontaneous speech, and to identify the features that differentiate the groups from each other. The task only takes a few minutes to administer, and poses a very limited demand on memory.

The described data provides a rich resource for future research in many areas of language impairment and has been constructed to facilitate analysis with natural language processing and corpus linguistics techniques since it can be analysed at all levels of linguistic structure, i.e., phonetic, lexical, syntactic, semantic and pragmatic. Whether the chosen picture description task will be suitable for detecting subtle language deficits in patients with subjective and mild cognitive impairment, is a question we are currently investigating. We anticipate that a single image might limit our future findings, however we believe that the breadth of analysis as outlined above and the combination of features of both the audio signal (e.g., prosody) and text (e.g., information content units) could provide a solid ground for analysis and experimentation. Moreover, we also intend to relate the performance of linguistic variables to available results of the neurodegenerative testing and other cognitive screening tools (e.g. MMSE) for the same population.

In parallel with the analysis of the available data, we also plan to extend this research by repeating the Cookie-theft experiment (and recordings) in the near future, that is in 18 months after the first recordings took place, and described in this paper. This might be useful for longitudinal assessment and comparison of the semi-spontaneous oral text production and for the evaluation of possible degradation, maintenance or recovery of skills. Moreover, we have also concrete ideas on how to complement the corpus with other types of oral tests we intend to record, such as semantic word fluency (category animals) and also spontaneous speech production/semi-structured conversation in which the participants will be encouraged to talk about a predefined task-oriented dialogue topic such as a map route (*cf.* Andersson et al., 1991). These additions are *already* approved by the local ethical committee review board (ref. T021-18). Finally, we also intend to acquire written descriptions of the Cookie-theft from previous Swedish studies. This will enable us to have a more balanced corpus of the Cookie-theft.

9. Acknowledgements

This work has received support from *Riksbankens Jubileumsfond* - The Swedish Foundation for Humanities and Social Sciences, through the grant agreement no: NHS 14-1761:1.

10. Bibliographical References

Ahmed, S., Haigh, A-M., de Jager, C.A., and Garrard, P. (2013). Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain*. 136 (12): 3727-3737.

Anderson, A. et al. (1991). The HCRC Map Task Corpus. *Language and Speech*, 34, pp. 351-366.

Ash, S., Evans, E., O'Shea, J., Powers, J., Boller, A., Weinberg, D., Haley, J., McMillan, C., Irwin, DJ., Rascovsky, K., and Grossman, M. (2013). Differentiating primary progressive aphasias in a brief sample of connected speech. *Neurology*. 23;81(4):329-36.

Becker, J.T., Boiler, F., Lopez, O.L., Saxton, J. and McGonigle, K.L.. (1994). The natural history of Alzheimer's disease. Description of study cohort and accuracy of diagnosis. *Arch Neurol*. 51(6):585-94.

Boersma, P., and Weenink, D. (2013). *Praat: doing phonetics by computer* [Computer program]. Version

6.0.19, retrieved in Aug. 2016 from <http://www.praat.org/>.

Boschi, V., Catricalà, E., Consonni, M., Chesi, C., Moro, A. and Cappa SF. (2017). Connected Speech in Neurodegenerative Language Disorders: A Review. *Front. Psychol.* 8:269. doi: 10.3389/fpsyg.2017.00269

Bottenberg, D., Lemme, M. and Hedberg, N. (1987). Effect of Story Content on Narrative Discourse of Aphasic Adults. *J Clinical Aphasiology*. 17. Pp. 202-209.

Bryant, L., Ferguson, A. and Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics*, 30:7, 489-518. DOI: 10.3109/02699206.2016.1145740

Bschor, T., Kühl, K.P., and Reischies, F.M. (2001). Spontaneous Speech of Patients with Dementia of the Alzheimer Type and Mild Cognitive Impairment. *Int Psychogeriatr.* Vol 13:3, pp. 289-298.

Catricalà, E. et al. (2017). SAND: a Screening for Aphasia in NeuroDegeneration. Development and normative data. *Neurol Sci*. 38:1469–1483. DOI 10.1007/s10072-017-3001-y

Chand, V., Baynes, K., Bonnici, L.M. and Tomaszewski Farias, S. (2010). *Analysis of Idea Density (AID): A Manual.* Technical report, U of California at Davis. Available from: <https://www.researchgate.net/publication/267362822_Analysis_of_Idea_Density_AID_A_Manual>

Choi, H. (2009). Performances in a Picture Description Task in Japanese Patients with Alzheimer's Disease and with Mild Cognitive Impairment. *Korean Academy of Speech-Language Pathology & Audiology*. Pp. 326-337.

Croisile, B., Ska, B., Brabant, M-J., Duchene, A., Lepage, Y., Aimard, G. and Trillet, M. (1996). Comparative Study of Oral and Written Picture Description in Patients with Alzheimer's Disease. *Brain Lang.* 53(1):1-19.

Cromnow, K., and Landberg, T. (2009). *Skriftliga beskrivningar av bilden "Kakstölden". Insamling av referensvärden från friska försökspersoner*. Master's thesis. Division of Speech and Language Pathology, Karolinska institute, Sweden. (In Swedish).

Cuetos, F., Arango-Lasprilla, J.A., Uribe, C., Valencia, C. and Lopera, F. (2007). Linguistic changes in verbal expression: A preclinical marker of Alzheimer's disease. *J of the International Neuropsychological Society*. 13, 433–439. Cambridge University Press.

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, *12*(3), 189-198.

Forbes, KE., Shanks, MF., and Venneri, A. (2004). The evolution of dygraphia in Alzheimer's disease. *Brain Research Bulletin.* 2004;63:19–2

Forbes-McKay, KE. and Venneri, A. (2005). Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. *Neurol. Sci*. 26, 243–254. 10.1007/s10072-005-0467-9

Fraser, KC., Meltzer, JA., and Rudzicz, F. (2016). Linguistic Features Identify Alzheimer's Disease in Narrative Speech. *J Alzheimers Dis.* 49(2):407-22. doi: 10.3233/JAD-150520.

Fraser, KC., Lundholm Fors, K, Eckerström, M., Themistokleous, C., and Kokkinakis, D. (2018). *Improving the sensitivity and specificity of MCI screening with linguistic information*. 2nd Resources and ProcessIng of linguistic, para-linguistic and extra-linguistic Data from people with various forms of cognitive/psychiatric impairments Workshop (RaPID-2, an LREC workshop). Miyazaki, Japan.

Gilles, E., Patterson, K., and Hodges, J. (1996). Performance on the Boston Cookie-theft picture description task in patients with early dementia of the Alzheimer's type: missing information. *Aphasiology*. 10:4:395-408.

Goodglass, H., Kaplan, E., and Barresi, B. (2001). *Boston Diagnostic Aphasia Examination* (3rd ed.). Philadelphia: Lippencott, Williams & Wilkins.

Groves-Wright, K., Neils-Strunjas, J., Burnett, R. and O'Neill, J.M. (2004). A comparison of verbal and written language in Alzheimer's disease. *J of Communication Disorders.* 37:2,, Pages 109-130

Hux, K., Wallace, S.E., Evans, K., and Snell, J. (2008). Performing Cookie-theft Picture Content Analyses to Delineate Cognitive Communication Impairments. *J of Medical Speech-Language Path*. Vol. 16:2. Pp. 83-99.

Jessen F. et al. (2010). Prediction of dementia by subjective memory impairment: effects of severity and temporal association with cognitive impairment. *Arch. Gen. Psychiatry*. 67(4). Pp. 414–422.

Johnson, S. C., et al. (2017). The Wisconsin Registry for Alzheimer's Prevention: a review of findings and current directions. *Alzheimers Dement (Amst).* 10:130-142. doi: 10.1016/j.dadm.2017.11.007.

Kavé, G. and Levy, Y. (2003). Morphology in Picture Descriptions Provided by Persons with Alzheimer's Disease. *J of Speech, Language, and Hearing Research.* Vol. 46, 341-352. doi:10.1044/1092-4388(2003/027)

Kavé, G. and Dassa, A. (2017). Severity of Alzheimer's disease and language features in picture descriptions. *Aphasiology*, DOI: 10.1080/02687038.2017.1303441

Kavé, G. and Goral, M. (2016). Word retrieval in picture descriptions produced by individuals with Alzheimer's disease. *J of Clinical and Experimental Neuropsychology*, 38:9, 958-966.

Kertesz, A. (1982). *Western Aphasia Battery*. NY Grune and Stratton.

Lind, M., Kristoffersen, K E, Moen, I., and Simonsen, HG. (2009). Semi-spontaneous oral text production: Measurements in clinical practice. *J of Clinical Ling & Phonetics*. 23:12, pp. 872-886. Taylor & Francis.

de Lira, J.O. et al. (2011). Microlinguistic aspects of the oral narrative in patients with Alzheimer's disease. *International Psychogeriatrics* Apr;23(3):404-12. doi: 10.1017/S1041610210001092.

de Lira, J.O., Minett, TSC., Bertolucci, PHF., and Ortiz, K.Z. (2014). Analysis of word number and content in discourse of patients with mild to moderate Alzheimer's disease. *Dement Neuropsychol.* Sept;8(3):260-265.

Lundholm Fors, K, Fraser, KC, and Kokkinakis, D. (2018). *Automated Syntactic Analysis of Language Abilities in Persons with Mild and Subjective Cognitive Impairment*. Proceedings of the 29th Medical Informatics Europe (MIE). Gothenburg, Sweden.

Machado Lima, T., Brandão, L., Pimenta Parente, M-A., and Peña-Casanova, J. (2014). Alzheimer's disease: cognition and picture-based narrative discourse. *CEFAC* vol.16:4 São Paulo. <http://dx.doi.org/10.1590/1982-021620147513>.

Mackenzie, C., Brady, M., Norrie, J., and Poedjianto, N. (2007). Picture description in neurologically normal

adults: Concepts and topic coherence. *Aphasiology*. Vol 21:3-4, pp. 340-354.

MacWhinney, B., Fromm, D., Forbes, M. P., and Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology*. 25 (11), 1286-1307.

Masrani V., Murray G., Field T.S. and Carenini G. (2017) Domain Adaptation for Detecting Mild Cognitive Impairment. In: Mouhoub M. and Langlais P. (eds) *Advances in Artificial Intelligence. AI 2017*. LNCS. Vol 10233. Springer.

Mueller, K., Koscik R., Hermann B., Johnson S. and Turkstra, L. (2017). Declines in Connected Language Are Associated with Very Early Mild Cognitive Impairment: Results from the Wisconsin Registry for Alzheimer's Prevention. *Frontiers in Aging Neuroscience*. 9: 437, pp. 1-14. Frontiers Media.

Nordlund, A. et al. (2005). The Goteborg MCI study: mild cognitive impairment is a heterogeneous condition. *J Neurol Neurosurg Psychiatry.* 76(11):1485-90.

Orimaye, S.O., Wong, JSM., Golden, K.J., Wong, C.P., and Soyiri, I.N. (2017). Predicting probable Alzheimer's disease using linguistic deficits and biomarkers. *BMC Bioinformatics*. 18:34

Pakhomov, SVS. et al. (2010). A computerized technique to assess language use patterns in patients with frontotemporal dementia. *J Neuroling.* 23(2):127–144.

Pekkala, S. et al. (2013). Lexical retrieval in discourse: An early indicator of Alzheimer's dementia. *Clinical Linguistics & Phonetics*. 27:12, 905-921.

Petersen, R.C. et al. (2001). Current concepts in mild cognitive impairment. *Arc of Neurology*, 58, 1985-1992

Rentoumi, V., Raoufian, L., Ahmed, S., de Jager, CA., and Garrard, P. (2014). Features and machine learning classification of connected speech samples from patients with autopsy proven Alzheimer's disease with and without additional vascular pathology. *J Alzheimer's Dis*. 42 Suppl 3:S3-17. doi: 10.3233/JAD-140555.

Ripich, D.N., Fritsch, T., Ziol, E. and Durand, E. (2000). Compensatory strategies in picture description across severity levels in Alzheimer's disease: A longitudinal study. *American J of Alzheimer's Disease*. Vol. 15:4.

Richardson, J.D., and Dalton, S.G. (2016). Main concepts for three different discourse tasks in a large non-clinical sample. *Aphasiology*. 30:1. Taylor and Francis.

Ritchie, K., and Touchon, J. (2010). Mild cognitive impairment: conceptual basis and current nosological status. *The Lancet*. Vol. 355:9199. Pp. 225–228. doi:10.1016/S0140-6736(99)06155-3.

Roark, B., Mitchell, M., Hosom, J-P., Hollingshead, K., and Kaye, J. (2011). *Spoken Language Derived Measures for Detecting Mild Cognitive Impairment*. IEEE Transactions of Audio Speech and Language Processing. 19(7): 2081–2090.

Seshadri, S., Beiser, A., Kelly-Hayes, M., Kase, CS., Au, R., Kannel, WB. and Wolf, PA. (2006). The lifetime risk of stroke: estimates from the Framingham Study. *Stroke*. 37(2):345-50.

Silveira, G., and Mansur, L. (2015). Analysis of prototypical narratives produced by aphasic individuals and cognitively healthy subjects. *Dement Neuropsychol*. 9(3):279-284.

Sirts, K., Piguet, O., and Johnson, M. (2017*). Idea density for predicting Alzheimer's disease from transcribed speech*. The 21st Conference on Computational Natural Language Learning (CoNLL), pages 322–332. Vancouver, Canada.

Tyche, O. (2001). *Subtila språkstörningar hos patienter med diagnosen MCI (Mild Cognitive Impairment) Del I: utifrån den tematiska bilden "Kakstölden"*. Master's thesis, Karolinska Institute, Sweden (In Swedish).

Wallin, A. et al. (2016). The Gothenburg MCI study: Design and distribution of Alzheimers disease and subcortical vascular disease diagnoses from baseline to 6-year follow-up. *J Cer Blood Flow Metab*. 36(1):114-131.

Williams, C., Thwaites, A., Buttery, P., Geertzen, J., Randall, B., Shafto, M., Devereux, B., and Tyler, L. (2010). *The Cambridge Cookie-Theft Corpus: A Corpus of Directed and Spontaneous Speech of Brain-Damaged Patients and Healthy Individuals*. The 7th Conference on Language Resources and Evaluation (LREC). Malta.

Yancheva, M., Fraser, K., and Rudzicz, F. (2015). *Using linguistic features longitudinally to predict clinical scores for Alzheimer's disease and related dementias*. 6th SLPAT. Pp. 134–139, Dresden, Germany.

Yancheva, M., and Rudzicz, F. (2016). *Vector-space topic models for detecting Alzheimer's disease*. 54th Annual Meeting of the Association for Computational Linguistics. Pp 2337–2346. Berlin, Germany. Association for Computational Linguistics.

Wechsler, D. (1997). *Wechsler Memory Scale - Third Edition Manual.* The Psychological Corporation.

Weissenbacher, D. et al. (2016). *Automatic Prediction of Linguistic Decline in Writings of Subjects with Degenerative Dementia.* NAACL-HLT. Pp 1198-1207. c 2016 Association for Computational Linguistics