

An acoustic and lexical analysis of emotional valence in spontaneous speech: Autobiographical memory recall in older adults

Deniece S. Nazareth^{1,2}, Ellen Tournier³, Sarah Leimkötter¹, Esther Janse³, Dirk Heylen², Gerben J. Westerhof¹, Khiet P. Truong²

¹Psychology, Health and Technology, University of Twente

²Human Media Interaction, University of Twente

³Centre for Language Studies, Radboud University Nijmegen

d.s.nazareth@utwente.nl

Abstract

Analyzing emotional valence in spontaneous speech remains complex and challenging. We present an acoustic and lexical analysis of emotional valence in spontaneous speech of older adults. Data was collected by recalling autobiographical memories through a word association task. Due to the complex and personal nature of memories, we propose a novel coding scheme for emotional valence. We explore acoustic properties of speech as well as the use of affective words to predict emotional valence expressed in autobiographical memories. Using mixed-effect regression modelling, we compared predictive models based on acoustic information only, lexical information only, or a combination of both. Results show that the combined model accounts for the highest proportion of explained variance, with the acoustic features accounting for a smaller share of the total variance than the lexical features. Several acoustic and lexical features predicted valence. As a first attempt at analyzing spontaneous emotional speech in older adults autobiographical memories, the study provides more insight in which acoustic features can be used to predict valence (automatically) in a more ecologically valid setting.

Index Terms: older adults, autobiographical memory recall, life events, sentiment analysis, valence, speech analysis

1. Introduction

With the growing aging population, there is an increasing need for emotion recognition technology that can help support not only healthy older adults, but also older adults who have cognitive impairment such as dementia. To contribute to this need, we collected speech data of older adults and investigated predictors of valence in speech. We asked older adults to share their emotional memories (i.e., sad and happy) using autobiographical memory recall. These emotional memories were coded with a scheme of life events and rated on a valence scale ranging from negative to positive. In this paper, we introduce this novel coding scheme of valence for emotional memories as none of the existing schemes cater for the complexity of emotional memories. Also, we investigate acoustic and lexical features as predictors of the valence coded in emotional memories.

Earlier studies show that there are still considerable knowledge gaps and underexplored topics in emotions recognition research that leave room for investigation. Many of the studies on the (automatic) analysis of emotional expression have been carried out with healthy (young) adults, and it is unknown whether the results of these studies generalize to older adults. While *speech perception* of emotional expression by older adults has been investigated previously (e.g., [1]), much less attention

has been paid to *speech production* and the realization of emotional expression by older adults. In addition, while arousal has clear and well-established correlations with specific acoustic features, much less knowledge is available for valence in relation to acoustics [2, 3], in particular in spontaneous settings. Collecting spontaneous emotional speech of older adults in an ecologically valid way is complicated. Further, annotating observed emotions is difficult, especially in spontaneous situations where emotional expression and interpretation depends on the personal background of the sender and the context.

The contribution of this paper is three-fold: 1) we explore acoustic variables that were previously found to be predictive of valence in older adults' spontaneous speech, 2) we present a novel spontaneous Dutch emotional speech database that is collected through autobiographical memory recall with older persons, and 3) we present a novel valence annotation scheme for emotional memories.

2. Related work

2.1. Acoustic correlates of Valence

In general, acoustic variables are found to be better predictors of arousal than of valence (e.g., [2, 4]). Associating emotional valence with acoustic features has yielded inconsistent results with mixed success [5, 6, 7]. Some studies did not find any value in acoustic cues for the prediction of valence (e.g., [8]). Laukka et al. (2005) [4] found that a positive valence indicated a lower F_0 mean, a larger range in F_0 and faster speech rate [4]. Other studies have associated positive valence with fast speech rate [9] or a larger F_0 variability [9, 10]. Schröder et al. (2001) [11] found that negative valence was associated with longer pauses and increased voice intensity. It appears that no distinct acoustic profile for valence has yet been determined therefore this study will examine the relation between valence and acoustic features.

2.2. Autobiographical Memories

Autobiographical memories can be defined as memories of an individual's life. These memories can be of past experiences, facts about themselves or encounters with other people [12, 13]. Autobiographical memory recall involves a complex construction of putting together mental representations of these past events by connecting a rich and emotional array of details and information to the related retrieved memory [14, 15]. Valence is an important dimension of autobiographical memory as it indicates to what extent a memory is seen as positive or negative [16, ?, 17]. Autobiographical memories can be structured by life scripts that represent a series of life events that take place in a specific order and represents a prototypical life course within a certain culture [18]. Research on life scripts generally includes

an overview of positive and negative life events categorized into mean prevalence, importance, and valence based on younger and older adults [19, 20, 21, 22, 23, 24]. Combining the findings of the research, an overview of life events can be constructed in terms of valence. These life events can then plausibly be used to provide structure in the complexity of an individual’s emotional autobiographical memory. In current research, no valence annotation scheme for emotional memories have been developed to the best of the authors’ knowledge. Therefore, these life events will be used as a basis for the novel valence annotation scheme for emotional memories in order to examine the relation of valence to acoustic and lexical features.

3. Materials and Methods

3.1. Participants

The current study is based on data from 11 participants (7 female; 4 male), aged between 65 and 85 years old ($M=73.5$; $SD=27.11$). Participants were recruited through advertisements in local newspapers. Participants had to be at least 65 years old, have normal or corrected vision and/or hearing and had to speak and read Dutch fluently. Exclusion criteria were memory problems, traumatic experiences and a pacemaker. Data collection was carried out by the first author. The interviews were conducted at the participant’s home or a location where the participant felt comfortable.

3.2. Experimental Design

3.2.1. Autobiographical Memory Test (AMT)

A revised version of the Autobiographical Memory Test (AMT) [25] was used as a word association task to elicit emotional memories. Participants were asked to recall three emotional memories for two cue words that differed in valence, namely sad and happy. Participants were instructed to concisely retrieve specific memories in their life that happened only once, on a certain time and day and did not last longer than a day. Two neutral cue words (grass; bread) were used to practice the retrieval of memories. The fixed order of the cue words was: grass, bread, sad (3x) and happy (3x).

3.2.2. Recording set-up

The recording setup of the interview included three microphones. A shotgun microphone was placed on the table in front of the participant and wireless lavalier microphones were used for the participant and interviewer. In this study, only the close-talk recordings of the participant were used.

3.3. Procedure

The study was approved by the Ethics Committee of the University of Twente (Nr 107426). Prior to the interview, participants signed the informed consent. The AMT was introduced and participants were asked to recall an emotional memory for the two neutral cue words (grass; bread) presented on a tablet. Then, participants were asked to retrieve three specific emotional memories for the cue words sad and happy (fixed order). Participants received a small gift for their participation. The interview lasted between 29 and 62 minutes ($M=45.02$).

3.4. Data

3.4.1. Valence of Emotional Memories scale (VEM)

We propose a novel coding scheme for establishing the valence of autobiographical memories called the Valence of Emotional

Memories scale (VEM). The VEM consists of a list of life events with corresponding averaged valence scores (ranging from 1=negative to 7=positive) that were compiled from prior research [20, 24, 21, 22, 23]. The VEM can be applied to transcripts. Examples of life events mentioned in this prior research are birth or death in which birth was assigned a higher valence score than death. Life events can be made more specific by adding a specific subject, for example, birth of a grandchild or death of a parent, which could affect the valence score. We observed that emotional memories are very personal and that the listed valence score associated with a certain life event sometimes did not reflect the correct valence of that specific memory. To allow for flexibility and to address these person-dependent experiences, a subjectivity score that could lower or increase the score by 1 point was introduced. As the compiled list of life events was not exhaustive, missing life events that we encountered in our data were added when needed (in total 3: “youth”, “hobby”, “death grandchild”). The valence scores of these added life events were established by selecting relevant multiple key words based on the research of Moors and colleagues [26] and by averaging these scores. In total, the valence scores of 47 life events were established. More information regarding the VEM can be found in [27].

3.4.2. Valence annotation with VEM

All cued memories were transcribed, anonymized and chunked into meaningful fragments based on related content. For example, a cued emotional memory about the loss of a loved one could be divided into several fragments within that memory such as serious disease, death and travelling. The emotional memories were chunked into fragments, establishing 207 fragments in total. Two raters separately evaluated the transcripts of the fragments to identify life events and to distinguish these from reflective fragments since the newly developed annotation scheme only applies to life events. Reflective fragments contain reflections about the participant’s own emotions which are not relevant for the current study and were hence discarded. Out of the 207 fragments, 183 were identified as life events (mean duration of 51.95s, std of 34.52s). Subsequently, the transcript of each life event was coded for valence by two raters according to the VEM scale. Inter-rater reliability was found to be $k = 0.64$, based on 207 fragments. Further consensus was achieved through discussion. Figure 1 depicts the top 10 of most coded life events in the emotional memories. Figure 2 shows the distribution of the VEM valence scores of 183 life events.

Figure 1: Top 10 occurrence of the VEM Life Events in the Emotional Memories with their corresponding valence scores.

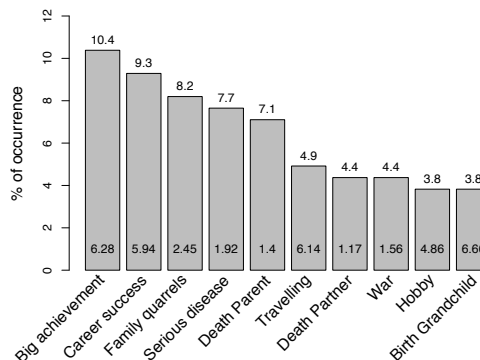
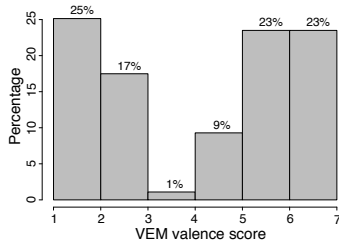


Figure 2: Distribution of the valence scores of the 183 life events ranging from 1 (negative valence) to 7 (positive valence).



3.5. Feature Extraction

3.5.1. Lexical features

Sentiment analysis belongs to a sub-field of natural language processing that aims at establishing the polarity of text. Polarity can be classified into negative, neutral or positive valence [28]. Sentiment analysis was applied to fragments by using the Pattern library [29]. Pattern is an open source lexicon consisting of Dutch words (mostly adjectives) with polarity strengths, additional intensity values and an algorithm that takes into account intensifiers, downtoners and negations. Downtoners apply when a sentiment of an adjective is strengthened or diminished by an adverb (e.g., “vreselijk mooi”, meaning “terribly beautiful”) whereas negations can distinguish between “niet blij”, “echt niet blij” and “niet echt blij”, meaning “not happy”, “really not happy” and “not really happy” respectively). As we are interested in the use of affective words in emotional memories, the Dutch lexicon was expanded with the lexicon of Dutch affective words based on [26]. Words with a polarity of $|p| < .03$ were excluded due to their neutral nature. Mean sentiment scores were calculated for each sentence within a fragment and averaged for a fragment. Sentiment scores ranged from -1 (negative) to 1 (positive) valence and were transformed into z-scores.

3.5.2. Acoustic features

Table 1 presents the acoustic features, that were expected to be associated with valence (based on previous research [30, 2]) and were minimally interrelated to each other (in our study). Acoustic features were extracted with Praat [31]. Mean, standard deviation and range of F_0 (Hz) and intensity (dB) were extracted, as well spectral balance (voice quality) and tempo information. The Hammarberg Index (dB) [32] indicates the energy distribution in the spectrum and is defined as the difference between the maximum energy in the lower frequency band (0–2000Hz) and maximum energy in the higher frequency band (2000–5000Hz). The articulation rate is extracted through a script by De Jong & Wempe [33] and is defined as the number of syllables per second without silence. Pause rate is defined as the number of silences per second. For each fragment, silent parts were identified by manually setting an intensity threshold in Praat (minimum silent duration 500ms, minimum sounding duration 150ms). Silent parts were discarded in F_0 , intensity, voice quality, and articulation rate feature extraction. All acoustic features were normalized per speaker through z-score transformation.

Table 1: Acoustic features used in the study.

F_0	mean, standard deviation, range
Intensity	mean, standard deviation, range
Voice quality	Hammarberg Index [32]
Tempo	mean pause duration, articulation rate, pause rate

3.6. Statistical analysis

In order to examine how acoustic features and lexical features predict valence in the emotional speech production of older adults, linear mixed-effects regression analyses were used to compare predictors of VEM valence scores (dependent variable) with participant as random intercept. To improve the model fit, predictors were step-wise excluded by removing the predictor with the highest non-significant p-value first. Based on the Akaike Information Criterion (AIC), the model with the most parsimonious fit was selected. The linear mixed-effects regression analyses (libraries lme4 [34] and lmerTest [35]) were computed with the statistical software R [36].

4. Results

4.1. Combining lexical and acoustic features

The combined model in which acoustic and lexical features acted as predictors was compared to models using acoustic-only and lexical-only features (Table 2). For the acoustic-only and lexical-only models, a step-wise exclusion was again performed to remove predictors that did not improve the model fit.

Table 2: Model comparison of the combined model, acoustic only and lexical only model.

	AIC	marginal R^2
Combined Model	672.74	0.460
Acoustic only	762.08	0.129
Sentiment only	686.01	0.384

Table 2 shows that out of the three models, the combined model explained most variance ($R^2 = 0.460$) whereas the acoustic features only ($R^2 = 0.129$) contributed to a smaller share of the total variance than the lexical features ($R^2 = 0.384$). The acoustic features turned out to better predict valence in the acoustic-only model than in the model that also contained the sentiment predictors: the R^2 is higher for the acoustics-only model than the difference in R^2 between the combined model and sentiment-only model (i.e., $0.460 - 0.384 = 0.076$).

4.2. Lexical features

Sentiment was found to be significant ($p < .001$), thus a higher mean of sentiment in a fragment was associated with a more positive VEM valence score (Table 3). So, the use of more affective words was related to more positive emotional memories.

Table 3: Estimates of the fixed effects of the best-fitting model of the VEM valence data; AIC = 672.745 with 183 observations ($df = 9$). Marginal $R^2 = 0.452$.

	β	SE	CI	p
Sentiment mean	1.19	0.108	0.98 - 1.40	<0.001
Duration pause mean	-0.29	0.120	-0.52 - 0.05	0.019
Pause rate	-0.40	0.121	-0.64 - 0.17	0.001
Hammarberg Index	-0.22	0.111	-0.44 - 0.00	0.050
F_0 range	0.30	0.138	0.03 - 0.57	0.030
F_0 mean	-0.28	0.136	-0.55 - 0.01	0.041

Note. Significance is indicated in bold face.

4.3. Acoustic features

The analysis of the acoustic features showed multiple significant predictors of the VEM valence scores (Table 3). Pause duration

also affected VEM valence score, such that a longer duration of pause was associated with lower VEM valence score. It appeared that older adults had longer pauses when discussing more negative emotional memories. Pause rate was found to be negatively associated with VEM valence scores. Thus, having more pauses in an emotional memory indicated a more negative valence for older adults. Regarding the Hammarberg Index, a marginally significant effect was found. It appeared that a lower Hammarberg Index correlated to a higher VEM valence score, suggesting that relatively more energy in the higher frequencies can be associated with higher valence and more vocal effort. The range of F_0 was associated with a more positive VEM valence score, meaning that a wider F_0 range was associated with more positively coded emotional memories. Finally, higher F_0 mean was associated with a more negative VEM valence score.

5. Discussion

One of the aims of this paper was to gain more insight which acoustic features contributed to the prediction of valence in older adults' spontaneous speech. Based on the results, multiple predictors of valence were found in terms of acoustic features and lexical features. When comparing the models containing either acoustic or sentiment features to the combined model containing both feature types, the explained variance of the two types of features did not simply add up, suggesting overlap in what was said and how it was said.

As expected, the choice of words (and their averaged valence score) is related to the valence of the event described. The use of more positively colored words in recalling emotional memories is related to more positively rated valence of those memories. This finding is not completely surprising since the memories were coded based on words only.

For the acoustic features, results showed that tempo and voice quality predicted the valence of life events. When older adults had longer and more pauses in their speech when discussing emotional memories (i.e., a slower tempo), it related to more negative life events. These findings are in line with previous research on valence in acoustics where negative valence was associated with longer pauses [11]. For voice quality, a low Hammarberg Index was marginally associated with higher valence. In other words, having relatively more energy in the higher frequency bands reflects more vocal effort and is known to relate to arousal [2]. The finding in Goudbeek & Scherer (2010) [2] that the Hammarberg Index is positively correlated with valence (steeper spectral slope is associated with more positive valence) is not replicated in our study. Instead, we found that a lower Hammarberg Index (a flatter spectral slope) was related to more positive valence. A possible explanation for this finding could lie in the nature of our data where we prompted "happy" and "sad" memories which differ not only in valence but also in arousal. As arousal or emotional intensity is not coded nor controlled, it is unknown to what extent the valence scores (or memories elicited) are confounded with arousal or emotional intensity. It is known that a flatter spectral slope can be related to high arousal [11, 37]. Hence, it could be that our finding of a lower Hammarberg Index that associated with more positive valence, is in fact a finding related to arousal rather than valence. Lastly, F_0 mean and range were found to predict valence in life events. The range of F_0 positively related to the valence of life events, meaning that older adults have a wider range of F_0 when discussing positive life events. This confirms the finding in other research that also found a positive association of valence with increased F_0 variability [10, 9]. With respect to F_0 , our results

are in line with previous research [9, 4] where a higher mean F_0 was found to be associated with more negative valence, whereas some studies did not find any significant relation [2]. A higher F_0 that is associated with lower valence could potentially be explained by the observation that, in our data, memories were sometimes expressed in a more emotionally intense way (e.g., crying). According to [4], emotional intensity shows parallels with high arousal that is strongly associated with high F_0 and other vocal cues. As the VEM life events were only annotated on transcripts, non-verbal vocalisations (e.g., crying) was not taken into account which brings us to the limitations of this study.

Possible important non-verbal cues of older adults when they spoke about their memories were not taken into account in this study and could have influenced the valence scores of life events. Although a first attempt was made with the VEM scale as a novel annotation scheme, it can be further developed into more clearly defined instructions which includes annotating on the combination of multi-modality (i.e., audio, video) to take non-verbal information into account. Another limitation of the study is the relatively small sample size ($N=11$). Although each participant had ± 16 life events, a relatively small sample size may have influenced the results. Lastly, the study focused only on the valence dimension as life events were based on valence scores from previous research. As a result, the discrimination of different emotional states is not possible at the moment [38, 39, 4]. As the findings in Hammarberg Index and F_0 show, it could be that emotional intensity or arousal act as a possible confounder. Therefore, emotional intensity or arousal should be taken into consideration when further developing the VEM coding scheme and should be included in future research.

For future research, we will also use the autobiographical memory recall in older adults with (mild) dementia as an emotion elicitation method. The present study was conducted in the context of a larger project that aims to investigate how emotional expressions can be (automatically) recognized using facial, vocal, and gestural expressions in people with dementia. The elicitation method of autobiographical memory recall was therefore partly chosen because the autobiographical memory abilities of older adults with dementia remain relatively intact [40]. Future research will compare the vocal expressions of valence in the autobiographical memories of healthy older adults and older adults with dementia, contributing to the need for emotion recognition technology for vulnerable groups.

6. Conclusions

Although challenging, a first attempt at analyzing spontaneous emotional speech in older adults' autobiographical memories was made. The study disambiguates which acoustic features can be used to predict valence in a more ecologically valid setting. The observed acoustic features are in line with prior research, therefore consolidating previous results on building a more complete profile of valence in spontaneous speech across the life span. A novel valence annotation scheme for emotional memories was introduced. The database of older adults' emotional memories will be made available to the research community as an example of a spontaneous-speech database.

7. Acknowledgements

The authors thank Michel-Pierre Jansen, Judith van Stegeren and Lorenzo Gatti for the expertise and assistance in this study. This research was funded by the Netherlands eScience Center as part of the NLeSC project 'Emotion Recognition in Dementia'.

8. References

- [1] J. Schmidt, E. Janse, and O. Scharenborg, "Perception of emotion in conversational speech by younger and older listeners," *Frontiers in psychology*, vol. 7, p. 781, 2016.
- [2] M. Goudbeek and K. Scherer, "Beyond arousal: Valence and potency/control cues in the vocal expression of emotion," *The Journal of the Acoustical Society of America*, vol. 128, no. 3, pp. 1322–1336, 2010.
- [3] I. B. Mauss and M. D. Robinson, "Measures of emotion: A review," *Cognition and emotion*, vol. 23, no. 2, pp. 209–237, 2009.
- [4] P. Laukka, P. Juslin, and R. Bresin, "A dimensional approach to vocal expression of emotion," *Cognition & Emotion*, vol. 19, no. 5, pp. 633–653, 2005.
- [5] J.-A. Bachorowski, "Vocal expression and perception of emotion," *Current directions in psychological science*, vol. 8, no. 2, pp. 53–57, 1999.
- [6] L. Leinonen, T. Hiltunen, I. Linnankoski, and M.-L. Laakso, "Expression of emotional-motivational connotations with a one-word utterance," *The Journal of the Acoustical society of America*, vol. 102, no. 3, pp. 1853–1863, 1997.
- [7] A. Protopapas and P. Lieberman, "Fundamental frequency of phonation and perceived emotional stress," *The Journal of the Acoustical Society of America*, vol. 101, no. 4, pp. 2267–2277, 1997.
- [8] C. Pereira, "Dimensions of emotional meaning in speech," in *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion*, 2000.
- [9] K. R. Scherer and J. S. Oshinsky, "Cue utilization in emotion attribution from auditory stimuli," *Motivation and emotion*, vol. 1, no. 4, pp. 331–346, 1977.
- [10] K. R. Scherer, "Acoustic concomitants of emotional dimensions: Judging affect from synthesized tone sequences." 1972.
- [11] M. Schröder, R. Cowie, E. Douglas-Cowie, M. Westerdijk, and S. Gielen, "Acoustic correlates of emotion dimensions in view of speech synthesis," in *Seventh European Conference on Speech Communication and Technology*, 2001.
- [12] M. Luchetti and A. R. Sutin, "Age differences in autobiographical memory across the adult lifespan: older adults report stronger phenomenology," *Memory*, vol. 26, no. 1, pp. 117–130, 2018.
- [13] R. Xu, J. Yang, C. Feng, H. Wu, R. Huang, Q. Yang, Z. Li, P. Xu, R. Gu, and Y.-j. Luo, "Time is nothing: emotional consistency of autobiographical memory and its neural basis," *Brain imaging and behavior*, pp. 1–14, 2018.
- [14] S. Sheldon, C. Fenerci, and L. Gurguryan, "A neurocognitive perspective on the forms and functions of autobiographical memory retrieval," *Frontiers in systems neuroscience*, vol. 13, 2019.
- [15] J. M. Talarico, K. S. LaBar, and D. C. Rubin, "Emotional intensity predicts autobiographical memory experience," *Memory & cognition*, vol. 32, no. 7, pp. 1118–1132, 2004.
- [16] P. J. Lang, "The motivational organization of emotion: Affect-reflex connections," *Emotions: Essays on emotion theory*, pp. 61–93, 1994.
- [17] D. Berntsen and D. C. Rubin, "Emotion and vantage point in autobiographical," *Cognition and Emotion*, vol. 20, no. 8, pp. 1193–1215, 2006.
- [18] D. Berntsen, M. Willert, and D. C. Rubin, "Splintered memories or vivid landmarks? qualities and organization of traumatic memories with and without ptsd," *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, vol. 17, no. 6, pp. 675–693, 2003.
- [19] D. C. Rubin, D. Berntsen, and M. Hutson, "The normative and the personal life: Individual differences in life scripts and life story events among usa and danish undergraduates," *Memory*, vol. 17, no. 1, pp. 54–68, 2009.
- [20] D. Berntsen and D. C. Rubin, "Cultural life scripts structure recall from autobiographical memory," *Memory & cognition*, vol. 32, no. 3, pp. 427–442, 2004.
- [21] A. Gryzman and S. Dimakis, "Later adults cultural life scripts of middle and later adulthood," *Aging, Neuropsychology, and Cognition*, vol. 25, no. 3, pp. 406–426, 2018.
- [22] S. M. Janssen and S. Haque, "Cultural life scripts in autobiographical memory," *Age*, vol. 71, p. 75, 2015.
- [23] S. M. Janssen and D. C. Rubin, "Age effects in cultural life scripts," *Applied Cognitive Psychology*, vol. 25, no. 2, pp. 291–298, 2011.
- [24] A. Erdoğan, B. Baran, B. Avlar, A. Ç. Taş, and A. I. Tekcan, "On the persistence of positive events in life scripts," *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, vol. 22, pp. 95–111, 2008.
- [25] J. M. Williams and K. Broadbent, "Autobiographical memory in suicide attempters," *Journal of abnormal psychology*, vol. 95, no. 2, p. 144, 1986.
- [26] A. Moors, J. De Houwer, D. Hermans, S. Wanmaker, K. Van Schie, A.-L. Van Harmelen, M. De Schryver, J. De Winne, and M. Brysbaert, "Norms of valence, arousal, dominance, and age of acquisition for 4,300 dutch words," *Behavior research methods*, vol. 45, no. 1, pp. 169–177, 2013.
- [27] D. S. Nazareth, M. P. Jansen, K. P. Truong, G. J. Westerhof, and D. Heylen, "Memoa: Introducing the multi-modal emotional memories of older adults database," in *ACII 2019 – 8th International Conference on Affective Computing & Intelligent Interaction*, in press.
- [28] S. M. Mohammad, "Sentiment analysis: Detecting valence, emotions, and other affectual states from text," in *Emotion measurement*, 2016, pp. 201–237.
- [29] T. D. Smedt and W. Daelemans, "Pattern for python," *Journal of Machine Learning Research*, vol. 13, no. Jun, pp. 2063–2067, 2012.
- [30] R. Banse and K. R. Scherer, "Acoustic profiles in vocal emotion expression," *Journal of personality and social psychology*, vol. 70, no. 3, p. 614, 1996.
- [31] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer. computer program, version 6.0. 50," 2019.
- [32] B. Hammarberg, B. Fritzell, J. Gaufin, J. Sundberg, and L. Wedin, "Perceptual and acoustic correlates of abnormal voice qualities," *Acta oto-laryngologica*, vol. 90, no. 1-6, pp. 441–451, 1980.
- [33] N. H. De Jong and T. Wempe, "Praat script to detect syllable nuclei and measure speech rate automatically," *Behavior research methods*, vol. 41, no. 2, pp. 385–390, 2009.
- [34] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.
- [35] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, pp. 1–26, 2017.
- [36] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2013. [Online]. Available: <http://www.R-project.org/>
- [37] K. R. Scherer, T. Johnstone, and G. Klasmeyer, "Vocal expression of emotion," in *Handbook of affective sciences*, 2003, pp. 433–456.
- [38] R. J. Larsen and E. Diener, "Promises and problems with the circumplex model of emotion." 1992.
- [39] R. S. Lazarus and R. S. Lazarus, *Emotion and adaptation*, 1991.
- [40] M. Kirk and D. Berntsen, "The life span distribution of autobiographical memory in alzheimers disease," *Neuropsychology*, vol. 32, no. 8, p. 906, 2018.