**Key literature:**

|  |  |  |  |
| --- | --- | --- | --- |
| Название | ГОСТ | Использованные метрики | Ключевая цитата |
| A Speech Recognition-based Solution for the Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech | Tóth L. et al. A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech //Current Alzheimer Research. – 2018. – Т. 15. – №. 2. – С. 130-138. | Speech rate, articulation rate, silent pause, hasitation ratio, pause-per-utterance ratio | significant differences for most of the acoustic parameters (speech tempo, articulation rate, silent pause, hesitation ratio, length of utterance, pause-per-utterance ratio) |
| Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech using ASR | Tóth L. et al. Automatic detection of mild cognitive impairment from spontaneous speech using ASR. – ISCA, 2015. |  | In this study, we sought to automate the feature extraction process by applying ASR. Unlike earlier authors, we used ASR to extract only a phonetic level segmen- tation and annotation. Furthermore, we took special care with filled pauses, which correspond to hesitations in most cases. We also extended the originally proposed features with further ones we considered informative. |
| Temporal parameters of spontaneous speech in Alzheimer's disease | Hoffmann I. et al. Temporal parameters of spontaneous speech in Alzheimer's disease //International journal of speech-language pathology. – 2010. – Т. 12. – №. 1. – С. 29-34. | articulation rate, speech tempo, hesitation ratio, and rate of grammatical errors | articulation rate, speech tempo, hesitation ratio, and rate of grammatical errors.  the clearest difference between the normal control group and the mild Alzheimer's disease group involved the hesitation ratio |
| Speech Analysis by Natural Language Processing Techniques: A Possible Tool for Very Early Detection of Cognitive Decline? | Asgari M., Kaye J., Dodge H. Predicting mild cognitive impairment from spontaneous spoken utterances //Alzheimer's & Dementia: Translational Research & Clinical Interventions. – 2017. – Т. 3. – №. 2. – С. 219-228. |  | a number of features regarding lexical, acoustic and syntactic aspects were significant in differentiating between mdMCI, eD, and CG (non-parametric statistical analysis). Some features, mainly in the acoustic domain also discriminated between CG and aMCI. |
| Towards the Speech Features of Mild Cognitive Impairment: Universal Evidence from Structured and Unstructured Connected Speech of Chinese | Wang T. et al. Towards the Speech Features of Mild Cognitive Impairment: Universal Evidence from Structured and Unstructured Connected Speech of Chinese //INTERSPEECH. – 2019. – С. 3880-3884. | filled pauses (e.g., “uh”, “mm”, “er”, “ah”, “zhege”, “nage” [“this” and “that” in Chinese]), disfluencies (e.g., false start, hesitation, stuttering) | Notably, acoustic parameters, describing quantitative speech and pauses in the signal (e.g., silence duration, speech duration, voice-silence ratio, verbal rate), were significantly modified due to the cognitive decline. |
| An internet-based screening test for Mild Cognitive Impairment (MCI) from spontaneous speech | Hoffmann I. et al. An internet-based screening test for Mild Cognitive Impairment (MCI) from spontaneous speech //Current Alzheimer Research. – 2015. | articulation rate, speech tempo, silent pause, hesitation ratio, length of utterance and break-to-utterance ratio | The higher hesitation ratio of the MCI group might reflect increased recall time. |
| Computer-Assisted Speech Analysis in Mild Cognitive Impairment and Alzheimer’s Disease: A Pilot Study from Shanghai, China | Qiao Y. et al. Computer-Assisted Speech Analysis in Mild Cognitive Impairment and Alzheimer’s Disease: A Pilot Study from Shanghai, China //Journal of Alzheimer's Disease. – 2020. – Т. 75. – №. 1. – С. 211-221. | percentage of silence duration, average duration of phrasal segments, average duration of silence segments, number of speech segments, number of long pauses, ratio of hesitation/speech counts and ratio of short pause/speech counts | All seven speech parameters significantly correlated with cognitive performance, with average duration of silence segments demonstrating the best correlation to cognitive performance on stepwise multiple linear regression analysis. |
| Szatloczki G. et al. Speaking in Alzheimer’s disease, is that an early sign? Importance of changes in language abilities in Alzheimer’s disease //Frontiers in aging neuroscience. – 2015. – Т. 7. – С. 195. | Szatloczki G. et al. Speaking in Alzheimer’s disease, is that an early sign? Importance of changes in language abilities in Alzheimer’s disease //Frontiers in aging neuroscience. – 2015. – Т. 7. – С. 195. | articulation rate, speech tempo, hesitation ratio, and grammatical error ratio | Results showed that articulation rate in mild and severe AD patients was significantly different from normal controls; furthermore, a difference among mild, moderate, and severe AD patients was also reported. Significant differences in speech tempo and hesitation ratio were found between all experimental groups, apart from moderate and severe AD patients, who performed similarly on both tasks. |
| Discourse Production and Normal Aging: Performance on Oral Picture Description Tasks | Cooper P. V. Discourse production and normal aging: Performance on oral picture description tasks //Journal of gerontology. – 1990. – Т. 45. – №. 5. – С. P210-P214. |  | No relationships between age and these measures were found, except for an increase in the use of prepositional phrases and indefinite words and longer pauses among older persons. Longer pauses may reflect cognitive slowing. |
| Speech Analysis by Natural Language Processing Techniques: A Possible Tool for Very Early Detection of Cognitive Decline? | Beltrami D. et al. Speech analysis by natural language processing techniques: a possible tool for very early detection of cognitive decline? //Frontiers in aging neuroscience. – 2018. – Т. 10. – С. 369. | e.g., silence segments duration, speech segments duration, verbal rate, transformed phonation rate, standardized phonation time, and standardized pause rate | Notably, these features seem to be able to distinguish, not only eD and mdMCI, but also aMCI from the CG. |
| Computer-based evaluation of Alzheimer’s disease and mild cognitive impairment patients during a picture description task | Hernández-Domínguez L. et al. Computer-based evaluation of Alzheimer’s disease and mild cognitive impairment patients during a picture description task //Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. – 2018. – Т. 10. – С. 260-268. |  | Phonetic variables were highly correlated with age and gender. |
| What happens when nothing happens? An investigation of pauses as a compensatory mechanism in early Alzheimer's disease | Pistono A. et al. What happens when nothing happens? An investigation of pauses as a compensatory mechanism in early Alzheimer's disease //Neuropsychologia. – 2019. – Т. 124. – С. 133-143. | No correlation was found between pause frequency and grey matter density. | Early AD patients use pauses differently according to the narrative type.  •  The frequency of pauses reflects different processes according to the narrative.  •  Patients use pauses as compensatory mechanisms in the earliest stages of AD. |
| Discourse Measures to Differentiate Between Mild Cognitive Impairment and Healthy Aging | Kim B. S., Kim Y. B., Kim H. H. Discourse measures to differentiate between mild cognitive impairment and healthy aging //Frontiers in aging neuroscience. – 2019. – Т. 11. – С. 221. |  | The proportions of dysfluencies and pauses were higher in both the aMCI participants and the naMCI participants than in the HC participants; however, there was no difference between the two subtypes of MCI. |
|  | Hawkins P. R. The syntactic location of hesitation pauses //Language and Speech. – 1971. – Т. 14. – №. 3. – С. 277-288. |  | Two-thirds of all the pauses, and three-quarters of the pause-time, was found to occur at boundaries between clauses. Pauses occasioned by lexical items occurred more frequently at a group boundary than within the group. |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

* В исследованиях одни из наиболее значимых метрик для различения МКС и КГ были акустические параметры, в частности разные виды пауз (silent, filled)

**Qusetions:**

How long of a silence is considered a pause?

* До 200 – короткая, после 1000 – длинная [Campione E., Véronis J. A large-scale multilingual study of silent pause duration //Speech prosody 2002, international conference. – 2002.]
* short (0.15 seconds), medium (0.50 seconds), and long   
  (1.50 seconds) pauses
* считается от 200 мс [https://www.researchgate.net/figure/Mean-Length-of-Silent-Pauses-200-msec-or-more-Before-Clauses-of-Different-Functional\_tbl3\_235633576]
* Goldmann-Eisler [14] advocated the use of a **minimum** **pause** **duration** of 250ms
* encountered as part of this study included 40ms [17], 150ms [7], 270ms [18], and 1000ms   
  [**10**]. **Given** the **variability** in **pause** **threshold** **values**, it is **not** **immediately** **evident** what is an   
  **appropriate** **value** for **minimum** **pause** duration.
* Основываясь на материале, я бы выбрала минимальную длину паузы 200 мс. У некоторых участников паузы редко длились больше 200 мс, при этом мне показалось, что примерная длина пауз каждого участника не сильно отличается друг от друга.

What pausing strategies are there?

Shoud we take into consideration “breath pauses” (not filled or silent) and how?

**Relevance (?актуальность):**

Most of the studies emphasize the need for a longitudinal research

Russian language

Inhomogenuous results for different studies

Differentiate not only MCI and AD or severe and mild MCI nut MCI and CG

Pausing strategies и синтаксические позиции для детекции МКС не так широко исследованы

**Hypotheses\*:**

1. Pause per utterance
2. Pause duration (mean)
3. Silence/speech ratio
4. Filled pause count
5. Filled pause rate
6. Disfluency count (относится ли к паузам?)
7. Disfluency rate (относится ли к паузам?)
8. Will syntactical positions of pauses differentiate in CG and MCI?
9. Will pausing strategies differentiate in CG and MCI?

\*Will MMSE score correlate with 1-7 pausal features

**Annotation:**

1. Annotate all pauses and durations
2. Annotate filled pauses (ааа, эээ, как его…)
3. Annotate disfluencies (false start – с… столом, stuttering, hesitation)

**План действий:**

1. Обзор литературы
2. Подбор релевантных метрик
3. Переразметка материала
   1. Фактическая точность
   2. Разметка всех пауз длинной более 200 (?) мс
   3. Перепроверка длины пауз
   4. Разметка disfluencies (по большей части не размечались практикантами) – заикание, false start, hesitation и т д
   5. Разметка filled pauses (&=em, &=um)
   6. Переразметка utterances
   7. Возможно предварительно быстро проанализировать, что у разметчиков сильно различается, а что не очень, чтобы понимать, на что обращать внимание.?
4. Подсчет метрик из гипотез
5. Идентификация pausing strategies (+литература)
6. Придумать, как разметить syntactical positions of pauses
7. Статистика, корреляции
8. Интерпретация результатов
9. \*Попытка автоматизации (на будущее)

**План работы:**

1. Введение
   1. Предмет работы: Что такое MCI и направление современной научной мысли (ранняя (автоматическая) детекция)
   2. Цель работы: Найти акустические (паузальные) характеристики спонтанной речи, которые смогут приблизить науку к автоматической детекции МКС
   3. Актуальность моей работы (на материалах русского языка, необходимость лонгитюдного исследования, перспективность изучения акустических (особенно паузальных) характеристик)
   4. Материалы (образцы спонтанной речи участников с мягким когнитивным снижением и без него), способы достижения (статистический анализ) и методики (разметка спонтанной речи в Clan). Выборка – обратившиеся в Клинику памяти.
2. Основная часть
   1. Выбор исследуемых характеристик (что такое паузы, какие они бывают, какие метрики с ними связаны)
   2. Соотносится ли количество пауз с МКС (pause/utt, pause duration, silence/speech)?
      1. Описание метрик и их значения
      2. Статистический анализ, выявление корреляций
      3. Интерпретация
   3. Filled pauses (можно объединить с b)
      1. Описание метрик и их значения
      2. Статистический анализ, выявление корреляций
   4. \*Disfluency (можно ли считать паузами?)
   5. Syntactical positions of pauses
      1. Какие могут быть синтаксические позиции пауз?
      2. Как я размечаю синтаксические позиции
      3. Анализ и интерпретация
   6. Pausing strategies
      1. Выявление стратегий расстановки пауз в образцах речи (+какие стратегии уже идентифицированы в литературе?)
      2. Анализ и интерпретация
3. Заключение
   1. Результат работы: какие метрики оказались эффективными, а какие нет?
   2. Интерпретация полученных результатов (Что отражают паузы? Когнитивное снижение как таковое, замедление в поиске слов, что-то еще?)
   3. Discussion (Дальнейшие направления работы? Что можно было учесть, но не учтено?)

**Pausing strategies (and syntactical positions)**

What pausing strategies are there:

* Variables relating

Hypothesis (updated):

1. How does pausing behaviour correlate with
2. Does the stage of cognitive decline impact the prevalent pausing strategy (which type of pause is prevalent)?
3. How scattered/accumulative strategy is related to cognitive decline? (scattered: shorter pauses but more, accumulated: lengthy pauses but less)
4. Will syntactical positions of pauses differentiate in CG and MCI? (
5. Will pausing strategies differentiate in CG and MCI?

\*Will MMSE score correlate with 1-7 pausal features

Measures:

1. Pause per utterance (number of pauses of 1 speaker/number of utt
2. Pause duration (mean)
3. Silence/speech ratio
4. Filled pause count
5. Filled pause rate
6. Disfluency count (относится ли к паузам?)
7. Disfluency rate (относится ли к паузам?)

Questions:

How to measure pausing distribution? I propose “pause accumulation rate: mean pause/n.pause”

Pause meta information:

“Я (0.123) [% change of subject; synt position]”

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Timepoint | Speaker | Pause length | Type of pause | Syntactical location | Syntactical location | Syntactical location |  |
| T1 | AB0123 | 0.234 | (0.223) empty pause | Inside clause | Inside utterance | Before NP |  |
| T2 |  |  | (0.223) &=smth filled pause | Between clauses | Outside utterance | Before AdjP |  |
| T3 |  |  | <а:>ананас [# 0.223]  lengthening |  |  | Before VP |  |
|  |  |  |  |  |  | Before AdvP |  |

|  |  |  |  |
| --- | --- | --- | --- |
| Название | ГОСТ | Использованные метрики | Ключевая цитата |
|  | Van Donzel M. E., Koopmans-van Beinum F. J. Pausing strategies in discourse in Dutch //Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP'96. – IEEE, 1996. – Т. 2. – С. 1029-1032. |  | Results show that there are basically three different types of pausing: **silent pauses, filled pauses, and lengthening of words**. Speakers apply these means in different ways to achieve pausing, by using one specific pause type or a combination of more than one. The way of applying pausing is rather uniform within one speaker, whereas the choice of a particular **strategy** is largely speaker dependent. |
| https://www.isca-speech.org/archive\_open/nolisp05/nol5\_042.html | Esposito A. Pausing strategies in children //ISCA Tutorial and Research Workshop (ITRW) on Non-Linear Speech Processing. – 2005. | empty and filled pauses, and pho- neme lengthening | Results show that each category plays a different role in the discourse organization, with short pauses and medium pauses preceding utterances containing new added information and long pauses identifying **changes in scene**, **time and event structures**, and the functional role of delimitating paragraphs. Moreover, different pause durations also depend on the amount of information contained in the utterances and signal the cognitive effort required to convey new, or given information. The present data may be relevant in assessing the system of rules that underlie pausing means, and identify predictive schemes of the speech temporal structure useful to improve text-to-speech synthesis systems. |
| https://www.rle.mit.edu/soundtosense/conference/pdfs/fulltext/Saturday%20Posters/SB-Esposito-STS.pdf | Esposito A., Marinaro M., Palombo G. Children speech pauses as markers of different discourse structures and utterance information content //Proceedings of the International Conference: From sound to sense. – 2004. – Т. 50. – С. 10-13. |  | Table 3 gives, for each child, **the number of words (Wds), clauses (Cls), filler conjunctions (Fcs), and paragraphs (Phs) marked by a pause (#Ps)** |
|  | Deme A., Markó A. Lengthenings and filled pauses in Hungarian adults’ and children’s speech. – 2013.  MLA |  | The results revealed different usage and appearance of lengthenings in the two age groups, therefore, differences in speech skills and strategies can be concluded. LEs and FPs differ mostly in their position in the speech session between the age groups, which has implications regarding different planning strategies of adults and children. |
|  | Pistono A. et al. Pauses during autobiographical discourse reflect episodic memory processes in early Alzheimer’s disease //Journal of Alzheimer's disease. – 2016. – Т. 50. – №. 3. – С. 687-698.  MLA |  | pausing behavior (**frequency, duration, and location**) |
|  | Kahng J. The effect of pause location on perceived fluency //Applied Psycholinguistics. – 2018. – Т. 39. – №. 3. – С. 569-591.  MLA |  | The findings suggest that listeners seem to be sensitive to pause location and to understand that pauses within clauses tend to **reflect reduced cognitive fluency**. |
|  | Pastoriza-Dominguez P. et al. Speech pause distribution as an early marker for Alzheimer's disease //medRxiv. – 2021. – С. 2020.12. 28.20248875. |  | Previous discussions addressing the use of these distribution type have proposed temporal thresholds differentiating short pauses from long speech pauses setting this barrier at 268ms [94], 323ms [120] or more recently at 338ms [93]. However, other authors have suggested up to three pause types: short (< 200ms), intermediate (200 − 1000ms) and long pauses (> 1000ms) [90]. The classification of “short” and “long” pauses was long reduced to a conceptual discussion about articulatory and/or respiratory (short) pauses on the one hand, and (long) cognitive pauses on the other [121,122]. |
|  | Angelopoulou G, Kasselimis D, Makrydakis G, Varkanitsa M, Roussos P, Goutsos D, Evdokimidis I, Potagas C (2018) Silent pauses in aphasia. Neuropsychologia 114, 41–49. |  | 338ms long-short pause barrier |
|  |  |  | Discourse organization: total number of words in the narrative, total number and mean length of pauses, total speech duration, speech rate and pause rate per 100 words;  Pause use: **between-utterance, within-utterance grammatical and within-utterance non-gram- matical pause** rates per 100 words; percentages of between-utterance pauses, within-utterance gram- matical pauses and within-utterance non- grammatical pauses. |
|  |  |  |  |
|  |  |  |  |

A characteristic of spontaneous speech, as well as of other types of speech, is the presence of silent intervals (empty pauses) and vocalizations (filled pauses) that do not have a lexical meaning. Such pauses seem to play a role in controlling the speech flow. Several studies have been conducted to investigate the system of rules that underlie speaker pausing strategies and their psychological bases. Research in this field had shown that pauses may play several communicative functions, such as build up tension or generate the listener’s expectations about the rest of the story, assist the listener in his task of understanding the speaker, signal anxiety, emphasis, syntactic complexity, degree of spontaneity, gender, and educational and socio-economical information (Bernestein, 1962; Goldman-Eisler, 1968; Abram & Bever, 1969; Kowal et al., 1975; Green, 1977; O’Connel & Kowal, 1983).

Several cognitive psychologists have suggested that pausing strategies reflect the complexity of neural information processing. Pauses will surface in the speech stream as the end product of a “planning” process that cannot be carried out during speech articulation and the amount and length of pausing reflects the cognitive effort related to lexical choices and semantic difficulties for generating new information (Goldman- Eisler, 1968; Butterworth, 1980; Chafe, 1987).