

**РАЗРАБОТКА МЕТОДИКИ СБОРА И АННОТИРОВАНИЯ
КОРПУСА ПОЛИЛОГОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ
МУЛЬТИМОДАЛЬНОГО РАСПОЗНАВАНИЯ ВАЛЕНТНОСТИ
ВЫСКАЗЫВАНИЙ И ДОМИНАНТНОСТИ ДИКТОРОВ**

Пищак А.В. (Университет ИТМО)

Научный руководитель – к.т.н. Махныткина О.В. (Университет ИТМО)

Исследования выполнены за счет финансирования университета ИТМО в рамках НИР №622281 «Разработка методов и алгоритмов для мультимодального распознавания валентности высказываний и доминантности дикторов в полилогах».

Введение. Речь имеет психосоциальный компонент, и каждое высказывание связано с эмоциональным состоянием говорящего. Фраза может быть эмоционально окрашена и иметь позитивную, негативную или нейтральную тональность. Полилог является одной из форм коммуникации говорящих. В процесс полилога каждый говорящий может быть вовлечен в разной степени. От качества высказывания зависит характер и направление разговора. Участники могут принимать доминирующую роль, подстраиваться или менять эти роли. Для того, чтобы отследить эти процессы, необходимо собрать и проанализировать корпус данных.

Основная часть. Цель данной работы — разработка методики сбора данных, проведение пилотного эксперимента, доработка неточностей и разработка методики аннотирования, а также сбор корпуса. После изучения датасетов с появляющимся лидером [1, 2], удалось выявить основные характеристики аудиовизуальных данных и их сбора. В исследовании появляющийся лидер и доминантность диктора в полилоге являются взаимозаменяемыми понятиями. Необходимо учитывать просодические характеристики говорящего, грамотно выделять отрезки времени и определить минимальную анализируемую единицу, а затем обрабатывать информацию, в том числе и акустическую [3, 4, 5, 6]. Для этого обратимся к коммуникативному подходу, который включает в себя метод дискурсивного анализа, методы анализа коммуникативной активности и содержательной активности коммуниканта [7, 8] и к основным характеристикам полилога [9, 10]. Предположительно, отличительными чертами полилога с появляющимся лидером являются качество реплик, изменения некоторых акустических характеристик, эмоциональная составляющая (эмоции и тональность), а также факт перебивания. Валентность высказывания можно отследить на основании датасета MELD [11]. Так, разрабатываются основные положения методики сбора данных – проблемная тема или вопрос для обсуждения, формат и правила, минимальное количество дикторов, техническое оснащение (платформа zoom) и методика аннотирования.

Выводы. Результатом данной работы является разработка методики разметки, проведение пилотного эксперимента и апробация предполагаемой методики аннотирования, а также ее корректировка для дальнейшего применения, сбора большего корпуса и анализа данных для мультимодального распознавания валентности высказываний и доминантности дикторов в полилогах.

Список использованных источников

1. Кодзасов С.В., Кривнова О.Ф. Общая фонетика. – М.: Рос. гос. гуманит. ун-т, 2001. – 591 с.
2. Кудрявцева А. Ю. Реализация коммуникативного доминирования и его признание в композиционном аспекте полилога в массмедийном жанре ток-шоу // Вестн. Том. гос. ун-та. 2016. №405. URL: <https://cyberleninka.ru/article/n/realizatsiya-kommunikativnogo-dominirovaniya-i-ego-priznanie-v-kompozitsionnom-aspekte-poliloga-v-massmediynom-zhanre-tok-shou> (дата обращения: 25.01.2023).
3. Норман Б. Ю. Полилог: речевая реальность и литературный прием // Жанры речи. 2016. №2 (14). URL: <https://cyberleninka.ru/article/n/polilog-recheyaya-realnost-i-literaturnyy-priem> (дата обращения: 25.01.2023).
4. Baranov, A.N. (2014) Aktivnost' uchastnika kommunikatsii: metody lingvisticheskogo analiza [Activity of the communication partner: methods of linguistic analysis]. In: Selegey, V.P. (ed.) Komp'yuternaya lingvistika i intellektual'nye tekhnologii [Computational Linguistics and Intellectual Technologies]. Moscow: RSUH.
5. Jayagopi D. B. et al. Modeling dominance in group conversations using nonverbal activity cues

- //IEEE Transactions on Audio, Speech, and Language Processing. – 2009. – Т. 17. – №. 3. – P. 501-513.
6. Poria, Soujanya. MELD: A Multimodal Multi-Party Dataset for Emotion Recognition in Conversations. / Hazarika, Devamanyu & Majumder, Navonil & Naik, Gautam & Cambria, Erik & Mihalcea, Rada. - 2018.
 7. Bryzgunova E. A. Prakticheskaya fonetika i intonaciya russkogo yazyka. – М. : Izd-vo MGU, 1963. – 308 s.
 8. Яковлева Э. Б. Полилог третья форма речи? // Известия ВУЗов. Поволжский регион. Гуманитарные науки. 2007. №1. URL: <https://cyberleninka.ru/article/n/polilog-tretya-forma-rechi> (дата обращения: 25.01.2023).
 9. Бондарко Л.В., Вербицкая Л.А, Гейльман Н.И. и др. Фонетика спонтанной речи // Под ред. Светозаровой Н.Д. - Л.: Изд-во Ленингр. ун-та, 1988. - 248 с.
 10. Sanchez-Cortes D. et al. Emergent leaders through looking and speaking: from audio-visual data to multimodal recognition //Journal on Multimodal User Interfaces. – 2013. – Т. 7. – P. 39-53.
 11. Вольская Н., Коваль А., Коваль С., Опарин И., Погарева Е., Скрябин П., Смирнова Н., Таланов А. Синтезатор русской речи по тексту нового поколения // Материалы XXI международной конференции Диалог-2005. М., 2005.