

Психологічні аспекти реалізації голосових інтеракцій машина-людина в антропотехнічних системах з високим рівнем автоматизації

Голосовий канал взаємодії «машина-людина» розглядається як інструмент ефективної реалізації потенціалу людини в умовах інтелектуальної кабіни. Викладено основний зміст поточних емпіричних досліджень, спрямованих на вивчення індивідуальних особливостей сприйняття та переробки людиною голосових повідомлень високої інформаційної ємності, а також їх впливу на виконання різних елементів операторської діяльності.

Постановка проблеми.

В умовах, коли завдяки автоматизації докорінно змінюється структура діяльності людини з управління технічними системами, на операторських робочих місцях починають переважати більш узагальнені, багатовимірні та високоінтелектуальні інформаційні потоки. Людина, яка керує сучасними рухомими об'єктами з притаманним їм високим рівнем автоматизації, в меншій мірі долучена до рутинних завдань операторської праці, а деякі з цих завдань у сучасних кабінах зберігаються лише з міркувань утримання людини в потоці процесів управління та створення якнайкращих умов для відчуття контексту, фізичного часу та розуміння динаміки подій. При цьому головним призначенням людини-оператора в усе більшій мірі вважається евристична діяльність, а також робота з такими категоріями особливих ситуацій, опрацювання яких засобами автоматики може виявитись проблематичним через високий рівень невизначеності та труднощі алгоритмізації.

Це накладає відбиток на підходи до організації систем інформування людини-оператора. З'явилися приклади науково-дослідних та дослідно-конструкторських робіт зі створення інтелектуальних систем розпізнавання небезпечних ситуацій, в тому числі пов'язаних із функціональним станом людини, з метою ефективної інформаційної підтримки людини. Ця підтримка знаходить реалізацію у двох різних аспектах: в аспекті інформування людини про розпізнані небезпечні ситуації [1], а також в аспекті ситуативного розвантаження людини та мінімізації потоків інформації відповідно до поточних пріоритетів діяльності [2].

Візуальне представлення інформації в сучасних авіаційних скляних кабінах стає дедалі більш синтетичним та спрямованим на формування цілісної картини, представленої у дружній для людського сприйняття спосіб. В той же час системи голосового інформування використовуються більш ортодоксально, ніж системи візуального відображення: досвід використання голосових інформаторів в людино-машинних інтерфейсах пов'язаний перш за все з системами невідкладного інформування людини про певний факт або параметр, а також з метою розвантаження потоків візуальної інформації.

Проблеми передачі інформації засобами живої мови отримали

грунтовне вивчення з точки зору психофізіологічних можливостей людини щодо розрізнення звуків та слів, а також нормування характеристик їх відтворення системами синтезу [3]. Наразі при використанні звичайних систем голосового сповіщення не виникає проблем сприйняття та розуміння змісту мовленнєвих повідомлень в передбаченому спектрі експлуатаційних ситуацій, так само як не є проблемою також забезпечення надійності систем голосового управління, в яких голосові команди надходять від людини до машини (як приклад можна навести систему голосового управління польотом, реалізовану в кабіні літака-винищувача 4-го покоління Eurofighter Typhoon), оскільки людина в системах голосового управління використовує команди, які мають окреслену смислову спрямованість і можуть бути стандартизовані подібно до стандартів професійної лексики взаємодії в льотних екіпажах та в їх комунікації з авіаційними диспетчерами та керівниками польотів.

Але реалізація голосового каналу в людино-машинних інтерфейсах може стосуватися принципово інших задач, ніж голосове управління машиною та голосове сповіщення людини. Зміст цих задач повинен впливати з трендів розвитку систем інформаційної підтримки людини на робочих місцях з високим рівнем автоматизації, а також обумовлюватися специфічними властивостями природної мови людини. Це дозволить уточнити поле питань, які потребують наукового опрацювання. Вочевидь, ці питання лежатимуть на перетині когнітивної психології, психології індивідуальних відмінностей, психолінгвістики та інших галузей науки.

Природна мова в інтеракціях машина-людина.

Природна мова має особливу інформаційну сміть та виняткові можливості оперування складними смисловими категоріями будь-якої спрямованості, в тому числі з високим рівнем узагальнення, актуалізація яких може виявитися необхідною у конкретній ситуації діяльності. Природна мова є також інструментом широкого спектру впливів на людину, оскільки сприйняття голосових повідомлень може викликати різноманітні когнітивні, поведінкові, а також емоційні ефекти, впливати на психічний стан завдяки і смислу і інтонаціям повідомлення, допомагати подоланню ілюзій та стереотипів, сприяти мобілізації. Властивості мови дозволяють використовувати її для впливу на метаструктури свідомості людини. Мовленнєвий формат – це формат не тільки повідомлень і команд, але й коментарів, оцінок, зауважень, міркувань, постановки запитань, і загалом це можливість побудови певного дискурсу, здатного сформувати або трансформувати психологічні установки сприйняття та оцінки ситуації, а також рішень щодо способів дій.

Реалізація подібних властивостей природної мови в інтеракціях машина-людина в умовах інтелектуальної кабіни з високим рівнем автоматизації в умовах сучасних технологічних можливостей виглядає дуже привабливо і вже розглядалася нами у контексті ідей створення інтелектуальної системи голосової підтримки [4], але наразі ці ідеї залишаються в більшій мірі питанням перспектив і експериментів. Показовим у цьому сенсі є висловлювання, що з позицій інженерної психології «теоретизування про сприйняття мовлення, прикрите всіма його

лінгвістичними та засвоєними функціями, рясніє пастками» [3, с. 195]. Але в сучасних умовах ця думка свідчить лише про те, що подібні пастки мають бути досліджені для їх уникнення.

Прикладом експериментального дослідження можливостей застосування живої мови у людино-машинній взаємодії є вивчення ефектів взаємодії астронавта зі спеціально створеною і доставленою на борт МКС системою штучного інтелекту, здатною на голосову комунікацію з проявами ознак емоцій [5,6]. Зазначається, що на дебютному етапі цього проекту робот не матиме усіх передбачених можливостей, а дослідження фокусуватимуться на ефектах соціальної взаємодії між астронавтом та емоційно-інтелектуальним помічником [6].

Загалом проблеми мовленнєвої комунікації мають глибоку історію психологічних та міждисциплінарних досліджень. В контексті всебічного вивчення процесів мовлення з позицій психологічної науки, увага приділялася широкому спектру питань у їх комплексі, починаючи з питань зв'язків мовленнєвої взаємодії з психічними процесами сприйняття, мислення та пам'яті, емоційними аспектами комунікації, спілкування в малих групах, і закінчуючи питаннями міждисциплінарного змісту, пов'язаними з імовірнісними аспектами мовного моделювання, проблемами коду мови, можливостями штучного мовлення, а також комунікацією в комп'ютеризованому просторі [7].

Проблема використання голосових каналів в людино-машинних інтеракціях полягає в тому, що голосове повідомлення складного змісту з багатьма смисловими аспектами може отримувати різні інтерпретації, а однакові за змістом повідомлення можуть бути представлені наборами різних словоформ. Таким чином, сприйняття і розуміння складної інформації, представленої в мовленнєвому форматі, може виявитись більш непередбачуваним та залежним від багатьох змінних.

Це вимагає обережності при реалізації голосових каналів інформування в тих видах діяльності, де важливо уникнути смислових спотворень. Для запобігання таким явищам можуть використовуватись кортежі з кількох повідомлень однієї смислової спрямованості. В комунікації між людьми в соціотехнічних системах в якості основного засобу подолання проблеми помилок сприйняття та хибних інтерпретацій голосових повідомлень використовується стандартизація фразеології.

Є підстави вважати, що дотримання стандартів фразеології може бути більш важливим в комунікації між людьми в операторських командах, ніж при реалізації голосових звернень з боку інтелектуальної системи, здатної не тільки розпізнавати сценарії розвитку подій, але й безперервно отримувати і обробляти об'єктивну інформацію про поточні показники активності людини-адресата, її загальний функціональний стан та окремі психофізіологічні реакції, особливості реагування на кожне з отриманих повідомлень, а також на кожну подію та в кожний момент часу. Прикладом реалізації моніторингу активності людини з метою поточної корекції процесу її інформування є адаптивний інтерфейс, відомий під назвою Cognitive Adaptiv Man-Machine Interface (CAMMI) [2]. Можливо, технології поточної оцінки стану та

характеристик активності людини є однією з передумов для реалізації таких голосових інтеракцій «машина-людина», які б не мали жорстких обмежень стандартної лексики та охоплювали увесь спектр імовірних ситуацій підтримки людини відповідно до функціоналу живої мови з притаманними їй психологічними характеристиками.

Але для оптимізації процесу інформаційної підтримки діяльності людини може бути потрібно не тільки здійснювати моніторинг динаміки її стану, як це реалізовано у системах, подібних до САММІ, але й враховувати індивідуально-типологічні особливості кожної конкретної людини, якій спрямовується повідомлення. Особливо цікавим це питання постає стосовно голосового каналу взаємодії. Мовленнєва модальність дозволяє варіювати різними характеристиками повідомлення зі збереженням його основного змісту, реалізовувати різні комунікативні стилі тощо. Звернення людини до людини зазвичай вибудовується з урахуванням певної сукупності уявлень про ту людину, якій адресоване повідомлення, тож чому в гібридних командах має бути не так само?

З технологічної точки зору, інформація про індивідуальні особливості людини може бути прописана в спеціально передбаченому персональному психометричному ідентифікаторі, із забезпеченням доступу до цієї інформації бортових інтелектуальних систем для їх оптимального самоналаштування на взаємодію з конкретною людиною [8]. З метою перевірити деякі припущення про залежність сприйняття вербальної інформації високої смислової ємності, наданої людині-оператору шляхом голосових повідомлень, від різних індивідуально-типологічних змінних, нами було розроблене та розпочате експериментальне дослідження, структура якого складається з двох етапів.

Перший етап нашого дослідження спрямований на вивчення індивідуально-психологічних чинників якості виконання елементів операторської діяльності та витривалості до її навантажень в умовах одночасного перебування людини в двох контурах інформаційної взаємодії: 1-й контур передбачає реагування за певним правилом на випадкові комбінації аудіовізуальних сигналів апаратного інтерфейсу, а 2-й контур має голосову модальність і передбачає сприйняття запитань в форматі живого мовлення, вчасного розуміння сутності почутих запитань, визначення з відповіддю на кожне з них на основі простих операцій вербально-логічного мислення та вчасне надання голосових відповідей, які засвідчують таким чином, що запитання дійсно були сприйняті. Тривалість роботи кожного досліджуваного складала 300 с, аудіозапис запитань відтворювався через гучномовець, пауза між запитаннями складала 6 с. Фіксувалася якість обробки сигналів апаратного інтерфейсу, а також здійснювався моніторинг рівня психофізіологічного навантаження за показником електричної активності шкіри (ЕАШ), який вимірювався екзосоматичним методом Фере, перетворювався у цифровий формат та кожні 0.5 секунди протоколювався у файл Excel за допомогою спеціального програмного забезпечення.

Після 3-хв. відпочинку кожен досліджуваний протягом наступних 300 с ще раз виконував завдання з обробки аудіовізуальних сигналів (під час цієї проби також здійснювалась реєстрація ЕАШ), але це завдання на цьому етапі

вже не поєднувалося з задачею осмислення потоку голосових повідомлень.

Відмінності показників якості обробки аудіовізуальних сигналів апаратурного інтерфейсу в 1-й і 2-й пробах, а також відмінності показників ЕАШ в 1-й і 2-й пробах ми розглядаємо в якості показників індивідуальної толерантності до навантажень сприйняття та розуміння голосової інформації високої смислової ємності в процесі фоновій операторській діяльності. Спостерігалися значні індивідуальні відмінності психофізіологічної толерантності до суміщення фонового завдання обробки апаратурних сигналів та завдання сприйняття і осмислення інформації в модальності живого мовлення, що можна бачити по відмінностям у співвідношенні показника А та показника В у різних осіб (рис.1).

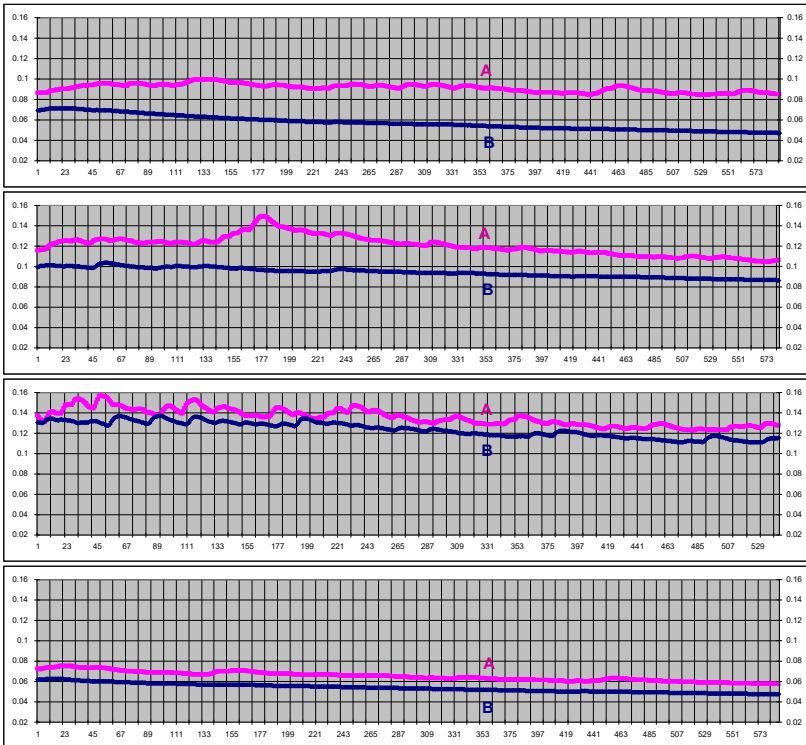


Рис. 1. Приклади співвідношення показників ЕАШ в першій (А) і другій (В) пробах

Після охоплення достатньої вибірки досліджуваних нами будуть проаналізовані статистичні зв'язки між показником індивідуальної толерантності до навантажень сприйняття складної голосової інформації та набором показників індивідуальних психофізіологічних відмінностей, отриманих на основі батареї тестів відповідно до завдань даного етапу.

Другий етап дослідження, за нашим задумом, полягатиме у виконанні досліджуваними наборів складних мисленневих задач, які мають певну зовнішню подібність але істотні сутнісні відмінності, що вимагає подолання когнітивних установок і ускладнює процес їх виконання. Цей етап передбачає дві проби, в першій з яких людині надходять голосові коментарі, які допомагають сформувати необхідний кут погляду на задачу, тоді як в другій пробі людина коментарів не отримує. Як і на першому етапі дослідження, обидві проби передбачають моніторинг рівня психофізіологічного навантаження за показником ЕАШ, а також роботу з батареєю тестів.

Висновок

Інтелектуальні системи голосової підтримки операторської діяльності можуть розглядатися як окремий клас інтелектуальних систем широкого спектру можливостей і потенційно спроможні набувати ознак віртуального суб'єкта в складі операторської команди. Функціонування таких систем може поширюватись на роботу з психічними станами людини та потребувати врахування її індивідуальних особливостей. Розробка інтелектуальних систем голосової підтримки вимагає досліджень на перетині когнітивної психології, психології індивідуальних відмінностей, психолінгвістики та ін. галузей науки.

Список літератури

1. Mouthaan Q., Ehlert P., Rothkrantz L. Situation recognition as a step to an intelligent situation-aware crew assistant system. // Proceedings of the 15th Belgium-Netherlands conference on Artificial Intelligence (BNAIC 2003), Nijmegen, The Netherlands. – P. 219-226.
2. Dorneich, M.C., Passinger, B., Beekhuyzen, M., Hamblin, C., Keinrath, C., Whitlow, S., Vašek, J. The Crew Workload Manager: An Open-loop Adaptive System Design for Next Generation Flight Decks. // Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Las Vegas, NV, Sept. 19-23, 2011. – P. 16-20.
3. Flanagan J. L., Allen J. B., Hasegawa-Johnson M. A. Speech Analysis Synthesis and Perception. Third Edition. – Springer-Verlag, New York, 2008. – 486 p.
4. Petrenko O. Possibilities of Using the On-Board Intelligent Voice Informing Systems in Complex Flight Situations. // Papers of the 18th International Symposium on Aviation Psychology (Wright State University, Dayton, Ohio, USA, 4-7 May 2015). – Red Hook, NY, 2015 (ISBN: 978-1-5108-0408-1). – P. 49-54.
5. Schmitz H.-Ch., Kurth F., Wilkinghoff K., Müllerschkowski U., Karrasch Ch., Schmid V. Towards Robust Speech Interfaces for the ISS. // Proceedings of the 25th International Conference on Intelligent User Interfaces Companion (IUI '20). Association for Computing Machinery, New York, NY, USA, 2020. – P. 110-111.
6. CIMON – the intelligent astronaut assistant. https://www.dlr.de/content/en/articles/news/2018/1/20180302_cimon-the-intelligent-astronaut-assistant_26307.html
7. Communication, Language and Meaning: Psychological Perspectives. Edited by George A. Miller – New York, Basic Books Inc., 1973. – 320 p.
8. Petrenko O. Man-machine symbiosis in aviation: new risks and capabilities in view of information technology expansion. // Papers of the 17th International Symposium on Aviation Psychology, (Wright State University, Dayton, Ohio, USA, 6-9 May 2013). – Red Hook, NY, 2013 (ISBN: 978-1-62993-579-9). – P. 116-121.