

DOI: <https://doi.org/10.32353/khrife.1.2020.24>

УДК 343.982.32

О. І. Брендель,

завідувач сектору досліджень відео-, звукозапису співробітник
лабораторії інженерно-технічних, екологічних, військових досліджень
та досліджень відео-, звукозапису ХНДІСЕ
ім. Засл. проф. М. С. Бокаріуса, м. Харків, Україна,
ORCID: <https://orcid.org/0000-0002-1723-0203>, e-mail: olgabrenda@ukr.net

ДЕЯКІ АСПЕКТИ ЗОРОВО-СЛУХОВОГО СПРИЙНЯТТЯ УСНОГО МОВЛЕННЯ В ЕКСПЕРТИЗИ ВІДЕО-, ЗВУКОЗАПISУ

Досліджено проблемне питання, яке все частіше постає в експертизі відео-, звукозапису, а саме: питання зорово-слухового сприйняття усного мовлення — визначення змісту розмови за її зображенням (читання з губ).

Метою статті є аналіз можливості та доцільності дослідження зорово-слухового сприйняття усного мовлення у межах проведення експертизи відео-, звукозапису, виокремлення особливостей такого дослідження. Розглянуто основні компоненти процесу читання з губ, можливості візуального дослідження зорово-слухової інформації з метою визначення змісту розмови. Звернуто увагу на особливості зорово-слухового сприйняття усного мовлення, проаналізовано найбільш інформативні фактори в загальній картині сприйняття усного мовлення за зображенням.

Проаналізовано можливості й доцільність дослідження зорово-слухового сприйняття усного мовлення у межах проведення експертизи відео-, звукозапису та виокремлено особливості такого дослідження.

Ключові слова: судова експертиза відео-, звукозапису; зорово-слухове сприйняття усного мовлення; читання з губ; відеозапис.

Постановка проблеми. У криміналістичних дослідженнях усе частіше експерти застосовують не лише звуко-, а й відеозаписи. Відеозаписи поєднують у собі два аспекти — звукову частину та відеозображення. Записи, здобуті за допомогою будь-яких відеокамер, є додатками до протоколів слідчих дій, оглядів місць подій, а також обшуків, слідчих експериментів, перевірки свідчень на місці тощо, їх експерти широко застосовують під час провадження як слідчих дій, експертних досліджень, так і під час проведення оперативно-розшукових заходів. За допомогою відеозаписів слідчі органи мають змогу більш повно сприймати та відображати події, проводити процесуальні дії, що мають значення у кримінальних провадженнях.

Під час аналізу відеозаписів, особливо якщо звукова частина супроводжується перешкодами, додаткову інформацію дає відеоканал, а за умови значного зашумлення запису звуку — візуальна інформація (артикуляція) може стати єдиним джерелом, за яким можна визначити зміст події, що відбувається.

Традиційно завданням судової експертизи відео-, звукозапису є ідентифікаційне дослідження диктора за фізичними і/або лінгвістичними ознаками голосу. Природно, що виникає питання про можливості використання візуальної інформації або як самостійного об'єкта дослідження (відеодоріжка відеозапису як джерело інформації про звукову частину зафіксованих подій), або як допоміжного, додаткового до слухового аналізу певного повідомлення. Під час розслідування у слідчих органів усе частіше виникає потреба у здобутті інформації щодо промовленого тексту лише за візуальною зорово-слуховою інформацією (читанням з губ) — визначенні змісту розмови.

Аналіз основних досліджень і публікацій. Експертиза відео-, звукозапису — це відносно молодий вид експертного дослідження. Теоретичні та практичні основи ідентифікації диктора за голосом закладено наприкінці 40-х років ХХ століття. У теорії та практиці цієї експертизи відбувається поступове створення її наукових засад, розробляються загальні й окремі методики дослідження. Інтенсивні дослідження з метою створення методик ідентифікації особи за голосом за фонограмами усного мовлення тривають і досі. Однак, певної методики проведення візуального дослідження зорово-слухової інформації (а саме: читання з губ) з метою визначити зміст розмови досі не створено.

Цю проблему також розглядало багато вчених-криміналістів іще від початку формування фоноскопичної експертизи як окремого виду експертизи. Можна сказати, що питання розпізнавання зорово-слухової інформації постало з моменту появи пристроїв відеозапису, які дають змогу багаторазово переглянути зафіксовані події.

Цю тематику порушували у своїх працях як вітчизняні, так і зарубіжні вчені, зокрема: Л. А. Чистович, А. В. Венцов ¹, Г. Фант ², Дж. Фланаган ³, В. С. Шупляков ⁴, Г. А. Куликов ⁵, DongSuk Yuk ⁶, Huang Xuedong ⁷ та ін.

¹ Чистович Л. А., Венцов А. В., Гранстрем М. П. и др. Физиология речи. Восприятие речи человеком/редкол.: А. В. Венцов, В. А. Кожевников, Д. М. Лисенко, Л. А. Чистович. Ленинград: Наука, 1976. 388 с. (Серия «Руководство по физиологии»).

² Фант Г. Акустическая теория речеобразования: пер. с англ. Москва: Наука, 1964. 284 с.

³ Фланаган Дж. Л. Анализ, синтез и восприятие речи: пер. с англ. А. А. Пирогова. Москва: Связь, 1968. 397 с.

⁴ Шупляков В. С. Глава 3. Физиология периферического отдела слуховой системы//Слуховая система/ред. Я. А. Альтман. Ленинград: Наука, 1990. 620 с. (Серия «Основы современной физиологии»).

⁵ Куликов Г.А. Нейрофизиологические основы обработки акустических сигналов//Акустика речи и слуха: сб. науч. работ. Ленинград: Наука, 1986. С. 56—87.

⁶ DongSuk Yuk. Robust speech recognition using neural networks and hidden Markov models. Adaptations using non-linear transformations. New Jersey: The State University of New Jersey, 1999.

⁷ Huang Xuedong. Spoken language processing: a guide to theory, algorithm and system development. New Jersey: Prentice Hall PTR, 2001. 910 p.

Внесок досліджень згаданих лінгвістів, фізіологів, психологів та інших учених полягає в розгляданні цієї проблеми загалом, не стосуючись питання можливості дослідження зображення відеоканалу як самостійного аспекту дослідження у криміналістичному аналізі мовлення на підставі видимої артикуляції.

Сучасний стан розвитку технологій дав змогу створити нейронну мережу, яка читає з губ з точністю понад 93 %. Розробники з Оксфордського університету створили першу у світі програму, яка розпізнає мовлення з губ на рівні цілих речень і робить це набагато краще за людину. Її робота заснована на використанні нейромереж і методів глибинного навчання. Наукову статтю дослідників, підготовлена до конференції ICLR 2017, оприлюднено у відкритому доступі.

Незважаючи на розроблення нових автоматизованих систем зчитування з губ, завдання зорово-слухового сприйняття усного мовлення в межах експертизи відео-, звукозапису сьогодні є актуальною та цікавою проблемою. Характерно, що кількість наукових робіт, присвячених невербальним засобам комунікації, зоровому сприйняттю усного мовлення, навчанню читання з губ значна, але цю прогалину в криміналістичному дослідженні відеозапису досі не заповнено.

Станом на сьогодні в методиках дослідження відео-, звукозапису відсутні як загальні підходи з дослідження зорово-слухового сприйняття усного мовлення, так і окремі напрями/шляхи, за якими можна проводити подібні дослідження.

Основною **метою статті** є аналіз можливостей використання візуальної інформації або як самостійного об'єкта дослідження (читання з губ), або як допоміжного, додаткового до слухового аналізу певного повідомлення та доцільність дослідження зорово-слухового сприйняття усного мовлення у межах проведення експертизи відео-, звукозапису, виокремлення особливостей такого дослідження, — візуального дослідження зорово-слухової інформації з метою визначення змісту розмови тощо.

Викладення основного матеріалу дослідження. Читання з губ — це зорове сприйняття усного мовлення за видимими артикуляційними рухами мовця.

Відомо, що під час читання з губ активується не лише зорова кора мозку, а й слухова, що також відбувається під час розпізнавання мовлення, яке ми чуємо. Навіть новонароджені діти намагаються імітувати губами і язиком рухи, які вони бачать у дорослих. Немовлята довше дивляться на джерело візуального стимулу, якщо йому відповідає і одночасно відбувається аудіальний стимул (на відміну від, наприклад, запису звуку). Діти у 4—8 місяців звертають особливу увагу на рухи губ тих, хто говорить як рідною для себе мовою, так і іншими мовами. Після 12 місяців вони вже не стежать за губами тих, хто говорить рідною мовою. У дорослому віці глухі читають з губ краще, аніж ті, хто чуєть, — завдяки постійній

практиці, різниця починає бути помітною з 14 років. Поміж глухих ця навичка краще розвинена у тих, хто втрачав слух поступово, а не раптово. Здебільшого читання з губ застосовують глухі та ті, хто погано чує. Успішність читання з губ залежить від здатності до образного мислення й аналітичного сприйняття, а також зорової пам'яті. Читання з губ ускладнюється через погане освітлення і наявність у мовця бороди або вусів ¹.

Також відомо, що сприйняття є результатом полімодальної діяльності, яка, передусім, має розгорнутий характер і лише згодом стає згорнутою, у якій образ сприйняття виникає на основі однієї модальності, а інші види чуттєвості лише допомагають відображенню. Найбільшою ємкістю сприйняття, стійкістю образу, як відомо, характеризується зоровий аналізатор, який може забезпечити надійну опору під час сприйняття та відтворення мовлення.

Читання з губ є надзвичайно складним процесом, у якому можна виокремити такі компоненти:

- зорове сприйняття артикуляційних рухів;
- кінстетичні імпульси, які надходять до кори головного мозку, які, відповідно до вчення І. П. Павлова, являють собою основний компонент мовлення та пов'язані з відображенням повторення мовленевих рухів;
- осмислення зорового сприйняття мовлення.

Вихідним моментом під час читання з губ є зорове сприйняття видимих артикуляційних рухів. За наявних візуального й акустичного каналів зв'язку, рухи артикуляційних органів не звертають на себе уваги, адже нормальна швидкість цих рухів переважає можливості нетренованого ока їх уловлювати. Багато вчених довели, що навіть слухове сприйняття значно підвищується у випадках, коли зображення мовця гарно проглядається. Для пояснення зорового сприйняття недостатньо знати лише фізіологічні закони процесів сприйняття образів оком, оскільки орган зору працює не ізольовано, а у складній динамічній системі разом з практикою, мисленням та всім попереднім досвідом особи.

У сприятливих умовах людське вухо вільно диференціює лише 42 звуки нашої фонетичної абетки. Око «виловлює» лише голосні та ті приголосні, які артикульовані губами. Пом'якшення, йотація, дзвінкість, глухість, як і наголос, оптичним шляхом практично не сприймаються, а якщо сприймаються, то надто слабко. Тому під час навчання читання з губ основну увагу приділяють звукам, які більш-менш чітко артикуються губами. Ці звуки є основними сигнальними ознаками, які дають змогу сприймати усне мовлення за допомогою зору. Чергування в процесі мовлення артикуляційних образів звуків, які видно, та пауз (які замішують звуки, яких не видно),

¹ Вікіпедія. Читання з губ. URL: https://uk.wikipedia.org/wiki/%D0%A7%D0%B8%D1%82%D0%B0%D0%BD%D0%BD%D1%8F_%D0%B7_%D0%B3%D1%83%D0%B1 (дата звернення: 14.05.2020).

складає своєрідний оптичний малюнок, який чітко сприймає око, а це значною мірою сприяє правильності читання з губ. У цьому разі сприймаються лише видимі ознаки артикуляційного складу — загальний склад губ, язика (не завжди), м'язів обличчя ¹.

Як відомо, і про це йдеться у багатьох роботах вітчизняних і зарубіжних учених, у навчанні читання з губ велике значення має зорове сприйняття видимої артикуляції. Таке навчання має відбуватися на матеріалі цілих фраз, а не окремих звуків ².

Деякі психологи зазначають зв'язок задовільних навичок читання з губ з такими факторами:

- здатністю до абстрактного мислення й образної уяви;
- зоровою пам'яттю;
- здатністю до аналітичного сприйняття ³.

Експерименти, які провадили в цій галузі, свідчать про те, що найбільшу інформативність у загальній картині сприйняття міміки обличчя мають рухи губ, зумовлені рухом нижньої щелепи. Читання з губ є цілісним процесом, у якому артикуляційні рухи створюють оптичні комплекси слів і найчастіше використовуваних словосполучень. Візуалізація надає більші переваги під час сприйняття емоційного боку мовлення, міміки обличчя, а також жестів, які сприяють інформативності візуального каналу зв'язку. Важливу роль у розпізнаванні відіграє наявність насамперед відомої ситуації спілкування.

Труднощі зорового сприйняття мовлення пов'язано насамперед з обмеженістю оральної (ротової) абетки, багатозначністю артикуляційних образів звуків, швидкістю звичайного розмовного мовлення та незнанням основ фонетики. Зрозуміло, що робота в аудіовізуальному та лише візуальному режимах суттєво відрізнятиметься, оскільки підключення зорового аналізатора до слухового значно поліпшує результати розпізнавання. Під час дослідження відеозаписів результати акустичного аналізу будуть кращими, ніж під час дослідження лише звукової частини або відеозображення окремо.

Спостереження за мовцем активує в мозку частини, відповідальні за слух, включно зі звивиною Хешла — ділянкою лівої півкулі мозку, де обробляється звукова інформація. Крім того, до нейронної мережі розпізнавання мовлення належать верхня скронева борозна й ділянка мозку, що відповідають за оброблення інформації про обличчя людей і відображення емоцій. Низка досліджень також відзначає активацію центру Брока — ділянку кори мозку, яка забезпечує аналіз пропріоцептивної інформації

¹ Потапов В. В. Некоторые аспекты зрительно-слухового распознавания речи// *Возможности судебной видеофонографической экспертизы*: сб. науч. тр./под ред. Н. П. Майлис, Э. К. Ребгун. Москва: ВНИИСЭ, 1989. С. 80—91.

² Untley R. Factors involved in the teaching and testing of lip reading ability through the use of motion pictures//*The Volta Review*, 1959.

³ Там само.

і являє собою кінетико-моторний аналізатор мовлення. По суті, цей центр забезпечує моторику мовлення. Водночас, як свідчать дослідження, читання з губ — це не просто сума аудіо-, візуальних сигналів. Досліджуючи цю тему, учені відкрили ефект Мак-Гурка — мультимодальність сприйняття людиною мовлення. Тобто Мак-Гурк зі своїми колегами свідчить, що наш мозок розпізнає мовлення, збираючи й обробляючи інформацію з різних органів чуття.

Під час прослуховування звуку (без візуального сигналу) людина розпізнає звук правильно. Під час прослуховування цього самого звуку з візуальним зображенням іншого звуку, який «накладений» на «вихідний» звук, людина розпізнає звук неправильно, звук не відповідає «вихідному» та не відповідає «накладеному» зображенню, учасники чують третій, неіснуючий склад. Цим ефект Мак-Гурка відрізняється від оптичних ілюзій — якщо про них розповісти піддослідному, він починає бачити їх такими, якими вони є насправді.

Вплив візуального образу на сприйняття аудіоданих виявляється навіть за умови явної невідповідності статі людини, яку записано на відео- й аудіо-запису. Якщо піддослідним показували чоловіка, який «говорив» жіночим голосом, аудіо і відео все одно давали інший (третій) за звучанням склад. Ефект Мак-Гурка формується в результаті аудіовізуальної інтеграції даних, що надходять до людського мозку. Учені зазначають, що це автоматична реакція головного мозку, не контролювана ззовні.

Навіть досвідчені особи, які вміють читати з губ, розуміють до 60 % інформації, оскільки кожна людина по-різному рухає губами. Хтось вимовляє текст, майже не відкриваючи рота, тому зрозуміти його важко. Краще, якщо людина активно артикулює, а її міміка висловлює емоції. Прикладом гарної артикуляції є розмова людей, які намагаються перекричати шум. Це активна артикуляція, вона важлива для сприйняття даних. Дуже складно розпізнавати мовлення чоловіка з вусами. Зазвичай навіть добре навчені люди, що мають досвід у цьому, не здатні довго читати з губ. Навіть серед людей, які погано чують, мало тих, хто вміє читати з губ досконало. Лише дехто здатний читати в будь-якому місці й у будь-який час, незалежно від артикуляції людини. Зазвичай глухі люди розуміють зміст розмови в межах 20—50 %.

Люди, які вміють читати з губ, фокусуються не лише на русі губ, точніше, вони фокусуються на всьому одразу, включно з рухами мовця, положенням зубів і навіть міміки, рухами головою, поглядом, жестикуляцією — будь-якими проявами емоцій.

Важливу роль у розпізнаванні усного мовлення за зображенням з губ відіграє і тривалість фрагмента мовлення. Стислих фрагментів інформації (або неповної частини розмови) може бути недостатньо для того, щоб сформувати «контекстне поле» та заповнити «прогалини» на місці незрозумілих слів. Повністю промовлене висловлювання зазвичай зчитується краще, аніж його частина.

Крім того, багато слів візуально схожі одне з одним (*бот, ром, гом, кот*). Зрозуміти слово можна, лише знаючи контекст розмови загалом. На відміну від аудіального сприйняття інформації, візуально, коли людина кричить, зрозуміти промовлене складніше, адже рот стає більшим, що ускладнює розпізнавання за зображенням. Навпаки, шепітне мовлення для візуального розпізнавання більш придатне.

Розробники з Оксфордського університету створили першу у світі програму *LipNet*, яка розпізнає мовлення з губ з точністю 93,4 % на рівні цілих речень і робить це набагато краще за людину. Однак, така точність — це все ж таки результат, здобутий у тепличних лабораторних умовах (людину показано крупним планом за відмінного освітлення, з високою роздільною здатністю, чіткістю тощо). Зрозуміло, що під час розпізнавання вільного мовлення людини результат буде набагато гіршим. Не кажучи вже про аналіз даних з реального відеозапису, де людину не фільмують крупним планом у необхідних умовах.

Висновки. Аналіз розглянутих аспектів зорово-слухового сприйняття усного мовлення (зокрема, читання з губ) свідчить про те, що підключення зорового аналізатора до слухового значно підвищує результати розпізнавання, порівняно з дослідженням звукової частини відеозапису окремо. Крім того, шепітне мовлення більш придатне для візуального розпізнавання змісту розмови за читанням з губ.

Щодо питання можливості використання візуальної інформації (читання з губ) як самостійного об'єкта дослідження у межах експертизи відео-, звукозапису з метою визначення змісту аудіоінформації, то воно має багато мінусів:

- мультимодальність сприйняття людиною мовлення;
- багато слів візуально схожі одне з одним, зрозуміти слово можна, лише знаючи зміст розмови загалом;
- тривалість фрагмента мовлення відіграє важливу роль у розпізнаванні усного мовлення за зображенням з губ;
- труднощі зорового сприйняття мовлення пов'язані з безліччю факторів, зокрема: обмеженістю оральної (ротової) абетки; многозначністю артикуляційних образів звуків; впливом швидкості звичайного розмовного мовлення на сприйняття інформації; різною артикуляцією та манерою мовлення у різних дикторів тощо;
- зміст розмови, здобутий читанням з губ, буде недослівним, він лише передаватиме суть розмови загалом, оскільки зазвичай «зчитується» 2—3 слова із 5 промовлених;
- крім кількісних елементів роздільної здатності, на якість зображення впливають чіткість відеофіксування, швидкість переміщення об'єктів у полі зору відеокамери, характеристики кольору та яскравості, ракурс знімання, наявність сторонніх «шумів» відеозображення тощо. Залежно від ситуаційних умов, усі визначені фактори впливають на якість зображення. Незадовільна якість

відеозображення суттєво напряму впливає на розпізнавання артикуляції. А зважаючи на наявність ефекту Мак-Гурка, результати дослідження незадовільного відеозображення та визначення змісту розмови за відеозображенням можуть викликати певні сумніви у достовірності «здобутого» за відеозображенням тексту.

Отже, в експертизі відео-, звукозапису використовувати навички читання з губ можна лише як допоміжного, додаткового до слухового аналізу певного повідомлення, що суттєво полегшить роботу експертів під час відтворення змісту розмови. Текстова відтворення змісту досліджуваної розмови «не є окремою експертною задачею, оскільки не потребує застосування спеціальних знань та може проводитись спеціалістом під час розв'язання питань персоніфікації мовленнєвого матеріалу чи застосуванні спеціальних методів покращення розбірливості усного мовлення»¹. Однак, текстове відтворення змісту розмови з персоніфікацією слів і висловлювань, що їх промовлено під час розмови, є обов'язковою частиною роботи експерта під час проведення ідентифікаційного дослідження з метою визначення тотожності мовленнєвої індивідуальності невідомого диктора, який брав у ній участь, і особи, яку перевіряють.

Зважаючи на наведене, використовувати візуальну інформацію як самостійний об'єкт дослідження, на думку автора, недоцільно, оскільки кількість і якість текстового відтворення, здобутого шляхом читання з губ, не гарантують точності здобутого результату, тому що вони залежать від безлічі факторів, зокрема, навичок особи, яка трактує зображення (її суб'єктивного сприйняття), та можуть свідчити лише про тему розмови (здобути дослівне повне текстове відтворення неможливо). Проте, використання навичок читання з губ як додаткового допоміжного інструмента під час складання змісту розмови в комплексі зі звичайним прослуховуванням, значно підвищить результати цієї роботи.

References

- Chistovich, L. A., Ventcov, A. V., Granstrem, M. P. i dr. (1976). *Fiziologija rechi. Vospriiatie rechi chelovekom*. Seriiа «Rukovodstvo po fiziologii». Moscow: Nauka [in Russian].
- DongSuk Yuk. (1999) *Robust speech recognition using neural networks and hidden Markov models*. Adaptations using non-linear transformations. – New Jersey: The State University of New Jersey.
- Fant, G. (1964). *Akusticheskaia teoriia recheobrazovaniia*. Moscow [in Russian].
- Flanagan Dzh. (1968). *Analiz, sintez i vospriiatie rechi* [in Russian].
- Huang Xuedong. (2001). *Spoken language processing: a guide to theory, algorithm and system development*. – New Jersey: Prentice Hall PTR.

¹ Про затвердження Інструкції про призначення та проведення судових експертиз та експертних досліджень та Науково-методичних рекомендацій з питань підготовки та призначення судових експертиз та експертних досліджень: наказ Мін'юсту України від 08.10.1998 р. № 53/5 (зі змінами та доповненнями). URL: <https://zakon.rada.gov.ua/laws/show/z0705-98> (дата звернення: 10.12.2018).

- Instruktsiia pro pryznachennia ta provedennia sudovykh ekspertyz ta ekspertnykh doslidzhen.* Naukovo-metodychni rekomendatsii z pytan pidhotovky ta pryznachennia sudovykh ekspertyz ta ekspertnykh doslidzhen: vid 08.10.98 № 53/5 (zi zminamy ta dop.). URL: zakon.rada.gov.ua/go/z0705-98 (data zvernennia: 10.12.2018) [in Ukrainian].
- Kulikov, G. A. (1986). *Neirofiziologicheskie osnovy obrabotki akusticheskikh signalov. Akustika rechi i slukha.* Sbornik trudov [in Russian].
- Potapov, V. V. (1989). *Nekotorye aspekty zritelno-slukhovogo raspoznaniia rechi. Vozmozhnosti sudebnoi videofonograficheskoi ekspertizy:* sb. nauch.tr. N. P. Mailis, E. K. Rebgun (Eds.). Moscow. VNII sudebnykh ekspertiz [in Russian].
- Shupliakov, V. S. (1986). *Fiziologiiia perifericheskogo otdela slukhovoii sistemy. Slukhovaia sistema.* Leningrad: Nauka [in Russian].
- Untley, R. (1959). *Factors involved in the teaching and testing of lip reading ability through the use of motion pictures.* Volta Revirw.
- Wikipediia. *Chtenie po gubam.* URL: https://ru.wikipedia.org/wiki/Chtenie_po_gubam (data zvernennya 14.05.2020) [in Russian].

О. И. Брендель

**НЕКОТОРЫЕ АСПЕКТЫ ЗРИТЕЛЬНО-СЛУХОВОГО
ВОСПРИЯТИЯ УСТНОЙ РЕЧИ В ЭКСПЕРТИЗЕ
ВИДЕО-, ЗВУКОЗАПИСИ**

Исследован проблемный вопрос, который всё чаще возникает в экспертизе видео-, звукозаписи (а именно: вопрос зрительно-слухового восприятия устной речи) — установление содержания разговора по его изображению (чтение по губам).

Целью статьи является анализ возможности и целесообразности исследования зрительно-слухового восприятия устной речи в рамках проведения экспертизы видео-, звукозаписи, рассмотрение особенностей такого исследования; возможность использования визуальной информации или в качестве самостоятельного объекта исследования (чтения по губам), или в качестве вспомогательного, дополнительного к слуховому анализу определённого сообщения. Рассмотрены основные компоненты процесса чтения по губам, возможности визуального исследования зрительно-слуховой информации с целью установления содержания разговора. Обращено внимание на особенности зрительно-слухового восприятия устной речи, проанализированы факторы, дающие наибольшую информативность в общей картине восприятия устной речи по изображению.

Рассмотрено влияние визуального образа на восприятие речи по изображению, в частности: активной артикуляции, мимики, движений головой, положения зубов, жестыкуляции и других. Кроме качества изображения, на восприятия устной речи по изображению также влияет и продолжительность фрагмента речи: полностью произнесённое выражение обычно считается лучше, чем его отдельные части. Также в статье обращено внимание на многозначность артикуляционных образов звуков. Рассмотрены особенности эффекта Мак-Гурка — феномена восприятия, который демонстрирует взаимодействие между слухом и зрением в восприятии речи.

Проанализированы возможности и целесообразность исследования зрительно-слухового восприятия устной речи в рамках проведения экспертизы видео-, звукозаписи и освещены особенности такого исследования.

Ключевые слова: судебная экспертиза видео-, звукозаписи, зрительно-слуховое восприятие устной речи, чтение по губам, видеозапись.

O. Brendel

SOME ASPECTS OF VISUAL-AUDITORY PERCEPTION OF ORAL SPEECH WHILE VIDEO AND SOUND RECORDINGS EXAMINATION

A problematic issue that frequently arises in the examination of video and audio recordings, namely the question of visual and auditory perception of oral speech – the establishment of the content of a conversation based on its image (lip reading) – is considered.

The article purpose is to analyze the possibility and feasibility of examining the visual-auditory perception of oral speech in the framework of the examination of video and sound recordings, considering the peculiarities of such research; the ability to use visual information either as an independent object of examination (lip reading), or as a supplementary, additional to auditory analysis of a particular message. The main components of the process of lip reading, the possibility of visual examination of visual and auditory information in order to establish the content of a conversation are considered. Attention is paid to the features of visual and auditory perception of oral speech, and the factors that contribute enormously to the informative nature of the overall picture of oral speech perception by an image are analyzed.

The influence of the visual image on the speech perception by an image is considered, such as active articulation, facial expressions, head movement, position of teeth, gestures, etc. In addition to the quality of the image, the duration of the speech fragment also affects the perception of oral speech by the image: a fully uttered expression is usually read better than its individual parts. The article also draws attention to the ambiguity of articulatory images of sounds. The features of the McGurk effect – a perception phenomenon that demonstrates the interaction between hearing and vision while the perception of speech – are considered.

The analysis of the possibility and feasibility of examining visual and auditory perception of oral speech within the framework of the examination of video and sound recordings is carried out, and the peculiarities of such research are highlighted.

Keywords: forensic examination of video and sound recordings, visual and auditory perception of oral speech, lip reading, video recording.

Надійшла до редколегії 09.06.2020

Брендель О. І. Деякі аспекти зорово-слухового сприйняття усного мовлення в експертизі відео-, звукозапису//*Теорія та практика судової експертизи і криміналістики*: зб. наук. пр./редкол.: О. М. Ключев, В. Ю. Шепітько та ін. Харків: Право, 2020. Вип. 21. С. 349—358. DOI: <https://doi.org/10.32353/khrife.1.2020.24>.