

Выражение и распознавание эмоциональных лиц как разные процессы

А. В. Сладкоштиева¹, А. С. Стародубцев¹, А. В. Петракова²

¹ СПбГУ, Санкт-Петербург

² ФНЦ ПМИ, Москва

e-mail: sladkoshtieva27@gmail.com

Аннотация. В эмпирическом исследовании трудно разделить процессы выражения и распознавания эмоций. Сторонники теории базовых эмоций и сторонники конструктивизма переносят результаты исследований распознавания на то, как работает выражение. Мы считаем, что их следует исследовать отдельно. Для независимого рассмотрения процесса выражения используется нейронная сеть teachable machine. Эффективность нейронной сети сравнивалась с эффективностью людей. Нейронная сеть не показала эффективности выше, чем у респондентов. Нейронная сеть и люди эффективнее справляются с распознаванием валентности и эмоциональных семейств, чем с категоризацией на уровне конкретных слов. Обсуждаются ограничения нейронной сети в распознавании всего диапазона вариаций прототипических выражений эмоций.

Ключевые слова: базовые эмоции, конструктивизм, распознавание эмоций, выражение эмоций

Концепция П. Экмана предполагает наличие базовых эмоций, которые имеют универсальное выражение лица в разных культурах и одинаково эффективно воспринимаются представителями разных культур (Ekman, 1973). В противоположность, концепция конструирования эмоций Л. Ф. Барретт отрицает наличие базовых эмоций, постулируя, что эмоции — это лишь понятия, служащие для когнитивной системы объяснительным конструктом о том, что происходит в окружающем мире и организме. Концепция не предполагает связи выражения лица и испытываемой эмоции, признавая только наличие негативного и позитивного ядерного аффекта (Barrett, 2011). Теория социального научения эмоциям предполагает возможность существования специфических эмоций, присущих только данной культуре. Проблема исследования эмоциональных лицевых экспрессий состоит в том, что в исследованиях выражения и восприятия прибегают к распространению выводов о результатах исследо-

вания распознавания эмоциональных экспрессий на объяснение процесса их выражения. В данной работе предлагается сделать акцент на том, что выражение и восприятие лицевой экспрессии — это разные процессы.

Мы считаем, что наличие разных слов в языке для обозначения эмоций являются лингвистической «ловушкой», которая может мешать определить базовую эмоцию, потому что слова-эмоции могут обозначать разные моменты процесса одной и той же эмоции, они могут быть связаны с целью, причинами, контекстом, присущим конкретной культуре (Oatley, Johnson-Laird, 2014). Эти факторы влияют на распознавание людьми мимики, что не дает прямого доступа к изучению выражения. Мы считаем, что социальные факторы и особенности языка распознающего больше влияют на распознавание, чем на выражении эмоции. Мы предполагали провести исследование с помощью нейронной сети, которая имеет свои особенности получения информации из изображения, не тождественные человеческому восприятию. Хотя напрямую изучать выражение нельзя (так как что-то или кто-то получает сигнал о мимике определенным образом в любом случае), мы считаем, что нейросеть, которая имеет свои особенности получения информации, не тождественные человеческому восприятию, сможет получить более точную информацию о выражении. Нейросеть опирается только на визуальные паттерны материала для классификации, что позволяет независимым способом исследовать выражение эмоций. Планировалось сравнить результаты, полученные нейросетью и людьми. Зависимость от опыта употребления слов родного языка не будет влиять на ответы нейросети.

Гипотезы: 1) У людей и нейросети будет наблюдаться более высокая правильность ответов, чем при случайном угадывании, начиная с такого уровня описания эмоции как эмоциональное семейство (в понятие семейство входит базовая эмоция и ее субэмоции), а также в случае валентности (исходя из теории П. Экмана, мы считаем, что на уровне четкой категоризации эмоции не будет сходных лицевых паттернов для каждого отдельного понятия в языке, одинаковые лицевые паттерны появляются только начиная с уровня эмоциональных семейств по П. Экману); 2) Люди будут менее точно категоризировать эмоции по сравнению с нейросетью в отношении эмоциональных семейств и валентности (исходя из сравнения особенностей нейросети и человека (нейросеть не зависит от лингвистических особенностей языка, настроения, не устает), мы полагаем, что нейросеть справится лучше человека, будучи заточенной именно на визуальные паттерны изображения).

Метод

В пилотной версии эксперимента приняли участие 32 человека. В нашем эксперименте использовался набор изображений, разработанный А. В. Петраковой совместно с коллегами (Петракова и др., 2024). Авторы-

создатели датасета изображений придерживались принципов экологической валидности при создании изображений. В эксперименте предъявлялось 248 цветных изображений эмоциональных лиц с 14 различными эмоциями в 3-х вариантах интенсивности. Эмоции были причислены к эмоциональным семействам и валентности, согласно П. Экману о «*emotional families*».

Мы не задавали конкретных областей на фотографии для распознавания нейросетью. Используется нейронная сеть *teachable machine*, обученная на датасете изображений *ImageNet*. Для обучения нейросети используется трансферное обучение. Модель обучалась на 717 изображениях датасета, которые не использовались в эксперименте. Используется гетерогенное обучение. Таким образом минимизируется вероятность негативного переноса. Нейросеть обучалась с параметрами в 40 эпох, размером пакета 16, скорости обучения 0,001. Критерием обучения нейросети была *Cross-Entropy Loss*. Нейросеть и респонденты в качестве ответа выдавали категорию на уровне отдельных эмоций.

В качестве платформы для демонстрирования фотографий респондентам использовались *Google*-формы. Перед экспериментом в инструкции испытуемому были показаны определения 14 эмоций (Барабанщиков, Суворова, 2020). Также в течение эксперимента перед испытуемым было напоминание с определениями эмоций. Ставилась задача категоризации эмоциональных лиц по 14 категориям. Каждая фотография предъявлялась на отдельной странице в новом разделе.

Ответ засчитывался как верный в случае семейств, если ответ, данный испытуемым или нейросетью, совпадал с семейством эмоции, действительно изображенной на фото, в случае валентности ответ считался правильным, если ответ, данный испытуемым или нейросетью, совпадал с валентностью эмоции, действительно изображенной на фото.

Результаты

С помощью одностороннего критерия Стьюдента были проведены сравнения вероятности случайного угадывания и доли правильных ответов на таких уровнях описания эмоции как точная категоризация и валентность. Не обнаружено значимых различий ($t = 1.0239$, $df = 13$, $p = .32$) между случайной вероятностью правильного ответа (.07) и долей правильных ответов нейросети ($M = 0.11$, $SD = 0.13$) при точной категоризации. Обнаружены значимые различия ($t = 5.3568$, $df = 13$, $p < .001$) между случайной вероятностью правильного ответа и долей правильных ответов людей ($M = 0.22$, $SD = 0.1$) при точной категоризации. Обнаружены значимые различия ($t = 2.2808$, $df = 13$, $p < .05$) между случайной вероятностью правильного ответа (.5) и долей правильных ответов нейросети для валентности ($M = 0.62$, $SD = 0.2$). Также обнаружены значимые разли-

чия ($t = 6.3094$, $df = 13$, $p < .001$) между случайной вероятностью правильного ответа и долей правильных ответов людей для валентности ($M = 0.73$, $SD = 0.14$).

Для семейств мы провели G -тест. Эффективность респондентов статистически значимо отличается от случайной ($G = 471$, $df = 13$, $p < .001$). Эффективность нейронной сети статистически значимо отличается от случайной ($G = 1142.5$, $df = 13$, $p < .001$). Респонденты эффективнее случайной доли категоризируют по семействам 13 из 14 эмоций, нейронная сеть — 8 из 14 эмоций.

Для сравнения корректности ответов по каждой эмоциональной категории, в зависимости от проходящего эксперимент (нейросеть или респонденты) был проведен однофакторный дисперсионный анализ ANOVA для двух независимых сравниваемых выборок. Тест показал значимые отличия между долей правильных ответов респондентов и нейронной сети в точной категоризации, $F(1, 26) = 5.11$, $p < .05$, $\eta^2 = .16$. Люди ($M = 0.22$, $SD = 0.1$) справляются с задачей категоризации эмоций статистически лучше, чем нейронная сеть ($M = 0.11$, $SD = 0.13$). Сравнение корректности ответов для изображений, правильно определенных в эмоциональное семейство, показало незначимые отличия между точностью ответов респондентов и нейронной сетью, $F(1, 26) = 1.95$, $p = .17$, $\eta^2 = .07$: нейронная сеть ($M = 0.38$, $SD = 0.23$), респонденты ($M = 0.5$, $SD = 0.24$). Сравнение корректности ответов по категориям валентности эмоций показало незначимые отличия между точностью ответов людей-респондентов и нейронной сетью, $F(1, 26) = 2.98$, $p = .1$, $\eta^2 = .1$: нейронная сеть ($M = 0.62$, $SD = 0.2$), респонденты ($M = 0.73$, $SD = 0.14$).

Обсуждение

И нейронная сеть, и респонденты успешно справляются с категоризацией эмоций на уровне семейств и валентности, что косвенным образом свидетельствует о том, что выражение лиц имеет соответствующие паттерны семейств и валентности. Респонденты также выше случайного распознают эмоции на уровне конкретных эмоций. Респонденты справляются лучше нейронной сети на уровне категоризации конкретных эмоций. Нет статистически значимых различий в эффективности на уровнях эмоциональных семейств и валентности. Новая теория базовых эмоций Scarantino может объяснить низкую эффективность на уровне конкретных эмоций в совокупности с высокой эффективностью на уровнях эмоционального семейства и валентности. Успешность человека, в сравнении с нейронной сетью, говорит о наличии определенных паттернов мимики для культурно специфических эмоций. Однако, сравнение с похожими исследованиями категоризации базовых эмоций говорит о малой специфичности выражения лица для небазовых эмоций.

Список литературы

1. Барабанщиков В. А., Суворова Е. В. Оценка эмоционального состояния человека по его видеоизображению // Экспериментальная психология. 2020. Т. 13. №. 4. С. 4–24.
2. Петракова А. В. и др. Опыт создания российской базы лиц, изображающих различные эмоции: первый этап // Психология. Журнал Высшей школы экономики. 2024. Т. 21. №. 2. С. 423–431.
3. Ekman P. Universal facial expressions in emotion // Studia Psychologica. 1973. V. 15. N. 2. P. 140–147.
4. Feldman Barrett L. Constructing emotion // Psihologijske teme. 2011. V. 20. N. 3. P. 359–380.
5. Oatley K., Johnson-Laird P. N. The communicative theory of emotions: Empirical tests, mental models, and implications for social interaction // Striving and feeling. Psychology Press, 2014. P. 363–393.

УДК 81

Активация конкурентных стимулов при восприятии устной речи на иностранном языке в контексте гипотезы критического периода

Д. Н. Соколова, М. В. Попова

МГЛУ, Москва

e-mail: darya_sokolova_2004@mail.ru

Аннотация. Настоящий план исследования посвящен изучению особенностей формирования фонетико-фонологической компетенции при изучении иностранного языка в контексте гипотезы критического периода. В качестве метода оценки когнитивных подтверждений гипотезы критического возраста мы предлагаем дизайн исследования, в ходе которого предполагается анализ скорости активации целевого стимула при помощи айтрекера при распознавании устной речи на иностранном языке у обучающихся, изучающих практическую фонетику иностранного языка. Исследование будет проводиться на материале немецкого языка при участии носителей русского языка, для которых немецкий язык является первым иностранным языком.

Ключевые слова: гипотеза критического периода, фонетика, айтрекер, немецкий язык, когнитивные стратегии.