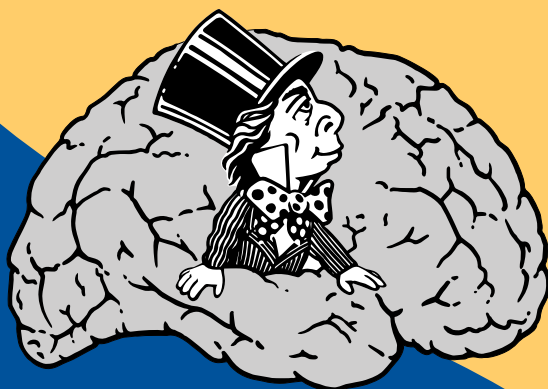


КОГНИТИВНАЯ НАУКА

В МОСКВЕ



НОВЫЕ ИССЛЕДОВАНИЯ

МАТЕРИАЛЫ
КОНФЕРЕНЦИИ
2019

Под ред. Е.В. Печенковой, М.В. Фаликман

УДК 159.9
ББК 88.25
К57

Когнитивная наука в Москве: новые исследования. Материалы конференции 19 июня 2019 г. Под ред. Е. В. Печенковой, М. В. Фаликман. – М.: ООО «Буки Веди», ИППиП. 2019 г. – 656 стр.

ISBN 978-5-4465-2346-7

УДК 159.9
ББК 88.25

ISBN 978-5-4465-2346-7

©Авторы статей, 2019

РУКИ, НОГИ, ГОЛОВА: ЗАДАЧИ, НАХОДКИ И ТРУДНОСТИ МУЛЬТИКАНАЛЬНОЙ АННОТАЦИИ

Ю. В. Николаева (1), А. А. Евдокимова (2)
julianikk@gmail.com, arochka@gmail.com

1 — МГУ имени М. В. Ломоносова, Москва; 2 — Институт языкознания
РАН, Москва

Аннотация. При изучении коммуникации в мультимодальной перспективе наблюдается тенденция к изоляции каналов: достаточно подробно изучаются жесты рук, существуют исследования по жестам головы; иногда какой-то из этих каналов сопоставляется с вербальным и/или просодическим. При этом про жесты ног удается найти критически мало работ, а комплексное изучение кинетических каналов (руки + голова + ноги + корпус + мимика) встречается только применительно к анимированным компьютерным агентам и в очень ограниченных контекстах. Наша работа представляет собой попытку сопоставить движения разных частей тела в рамках кинетической модальности, используя единый подход при аннотировании, учитывающий, однако, специфику каждого канала. В статье рассматривается такое явление, как мультиканальные кластеры — сочетание движений в разных каналах для выражения единого значения. Мы показываем, что общепринятая тенденция к изучению движений только одного канала приводит к игнорированию некоторых значимых жестов и к затруднениям при описании естественной коммуникации.

Ключевые слова: мультиканальная аннотация, жестикуляция, движения головы, корпус, движения ног, движения локтей, мультиканальные кластеры

Введение

При том что уже накоплены достаточно большие знания о мультимодальных средствах коммуникации (см., напр., Beattie, 2016; Litvinenko et al., 2018; Rehm, André, 2008), еще предстоит изучить, как взаимодействуют разные каналы, относящиеся к кинетической модальности. Среди невербальных кинетических каналов можно упомянуть мануальный (жесты рук), цефалический (движения головы), направление взгляда, мимический, движения ног и туловища. При этом надо отметить, что жесты рук, в которых движение кисти минимально по амплитуде и которые выполнены скорее локтем или плечом, требуют отдельного внимания — традиционно к мануальным жестам их не относят, так что для них, видимо, требуются отдельные каналы.

Данная статья посвящена сопоставлению движений в мануальном (движения рук) и цефалическом (движения головы) каналах и изучению их вза-

имодействия, а также особенностям движений ног и локтей как каналов в кинетической (то есть жестовой) модальности.

Методика

Материалом исследования послужил новый ресурс «Рассказы и разговоры о грушах»¹ (Кибрик, Федорова, 2018). Участникам записи предлагалось посмотреть и пересказать «Фильм о грушах» У. Чейфа (Chafe, 1980). В каждой сессии участвовало четыре человека: Рассказчик (N²), смотревший фильм, должен был передать его содержание Пересказчику (R), не видевшему фильма. При этом присутствовал Комментатор (C), который тоже видел фильм и после рассказа N мог добавлять и поправлять его, а R мог задавать им любые вопросы. После этих двух этапов (рассказ и обсуждение) наступал третий: к участникам присоединялся Слушатель (L), и R описывал ему содержание фильма.

Рассмотренный нами материал включал 3 записи (04, 22 и 23), в сумме продолжительностью 60 минут. Указанные записи состояли из видеозаписей трех участников, таким образом общая длительность видео составила 180 минут. Для всех участников были размечены движения кистей рук, локтей, головы и ног в программе ELAN.

Предлагаемые нами принципы аннотирования в ELAN предполагают описание движений в каждом канале отдельно и с выключенным звуком (На этом этапе разделяются движения и неподвижность, а для цефалического канала, то есть движений головы, — различаются собственные движения головы и такие, которые инициированы в другом канале; различие основано на степени включенности мышц шеи. Движение второго типа в канале головы описывается как движение-эхо либо перемещение). На втором этапе аннотирования, с учетом слов и просодии, а также действий других участников коммуникации, происходит разделение на коммуникативно значимые жесты и ориентированные на комфорт говорящего адапторы и смены позы. Таким образом, движения переосмыслились как фазы более сложно устроенных функциональных единиц — жестов или адапторов либо как смены положения. На этом этапе учитывались формальные характеристики движения (скорость, амплитуда, траектория и другие особенности движения) и функциональные признаки (воспринимается ли это движение как иллюстрация к речи; согласовано ли оно по времени, содержанию и функциям, которые оно выполняет в разговоре, со словами и просодией).

Уже на втором этапе проявляются особенности разных каналов. Если для мануальных жестов прототипическим является случай со вспомогательными фазами — подготовкой (когда рука переходит из положения покоя к той точке в жестовом пространстве, из которой начнется жест) и ретракцией (возвращением в положение покоя после завершения жеста), — то для других каналов (движения головы, локтей или ног) вспомогательные фазы могут отсутствовать или располагаться в другом слое, например в слое движений-эха. Так, для

1 Подробнее см. на сайте корпуса здесь: www.multidiscourse.ru

2 Здесь и далее указаны сокращения от англ.: N — narrator; R — reteller; C — commentator; L — listener.

кивка или отрицательного покачивания головой фаза подготовки может быть эхом от канала туловища или других каналов (например, движение туловища вперед или назад для кивка либо сдвиг вправо или влево для отрицательного покачивания); замах же назад, который можно заметить при некоторых видах особенно энергичных кивков, представляется потенциально значимым элементом: в английском языке, к примеру, различаются кивок вниз *nod* и кивок вверх *jerk*).

Для ног тоже разделялись функциональные типы движений: жесты, адапторы и смены положения. Ключевым для выделения жестов было согласование с речью по времени и семантическим признакам, вытекающим из формальных характеристик. Жесты ног еще мало изучены, однако данные нашего корпуса показывают, что, в зависимости от склонностей данного говорящего, их число может достигать от нескольких десятков до сотни на 10 минут записи. Примерами жестов ног может быть иллюстрация направления движения (вверх, вниз, в сторону) или взаимное перемещение референтов. Примером адаптора ног может служить почесывание ноги, когда человек для удобства придвигает ногу к руке.

При описании движений локтей рассматривались такие, амплитуда которых больше амплитуды смещения кисти на том же временном отрезке, при этом такое заметное движение локтя не является частью вспомогательной фазы для жеста руки.

Помимо общих для всех кинетических сигналов трех функциональных категорий — жестов, адапторов и смен поз — мы предлагаем смещения (для головы — когда изменение положения не связано с работой мышц, отвечающих за этот канал, в данном случае мышц шеи, а вызвано перемещением корпуса) и движения-эхо. Примером эха может быть отскок руки, быстро опустившейся на колени, или легкие покачивания головой при смехе.

Результаты

Сопоставляя движения в мануальном и цефалическом каналах (движения рук и головы), можно увидеть, что некоторые жесты и адапторы (от 0.2% до 15%, в зависимости от говорящего) задействуют два канала: например, потирая нос (адаптор в мануальном канале), человек наклоняет голову вниз (адаптор в цефалическом канале); показывая поля шляпы вокруг головы (жест в мануальном канале), говорящий поворачивает голову или делает движение вперед (жест в цефалическом канале). Анализ этих движений только в одном канале привел бы к тому, что целостную картину было бы невозможно описать.

Анализ взаимодействия мануального канала и канала ног выявляет следующие закономерности:

1) Поскольку все испытуемые в наших записях сидят, то положение покая рук иногда оказывается с опорой на ноги, и начало движения рук (выполнение жеста или адаптора) может сопровождаться эхом в ногах, изменением позы ног или отдельным жестом в ногах — ритмичными короткими движениями в ступнях или коленях (мы называем их «биты», от англ. *beats*). Традиционно такие движения относятся к жестам, поскольку обычно согласованы

с речью (McNeill, 1992), но их семантические значения непрозрачны, а кроме того, они иногда используются в функции адапторов и трактуются как выражение внутреннего состояния говорящего, не связанное с коммуникативным намерением. Биты могут быть отдельным предметом для исследования, так, у некоторых испытуемых (например, 22С) возникает синхронизация битов в кистях и в ступнях.

2) Классические моменты смены позы могут совпадать в разных кинетических каналах и собираться в итоге в одну позу туловища, а могут быть рассинхронизированы. При этом если руки занимали позицию на ногах в момент смены позы, то их движения можно рассматривать и как эховые в случае маленькой амплитуды и при отсутствии видимых усилий мышц, и как отдельный класс, близкий по своему значению к адапторам. Выделяется большой класс движений, которые по сути являются сложными изменениями позы (Litvinenko et al., 2018), собирающимися из совокупности смен положения в разных кинетических каналах; иногда смены позы в одном канале могут накладываться на жесты или адапторы в другом.

3) Следующий класс движений ног — это подчеркивание жестов рук, обычно с небольшим запаздыванием (встречается у 22N, 22С, 22R). Может выполняться опорной ступней, качанием ноги на другой ноге, обеими ногами (например, движение коленей к центру или от центра показывает движение двух персонажей относительно друг друга).

4) Ноги используются руками как материал для показа, как некая поверхность (например, 22N рисует руками на коленке, выполняя таким образом серию иллюстративных жестов).

Что касается движений локтя и мультиканальных кластеров с их участием, здесь тоже проявляются яркие индивидуальные различия (например, участник с кодом 04С опирается локтями на ноги, для него частое положение покоя — с опорой локтями на бедра, при этом локти иногда соскальзывают, и возврат в положение покоя задействует движения ног, локтей и корпуса).

В канале ног при опоре локтей на них у некоторых испытуемых отмечается отзеркаливание движений одного канала в другом. Так, например, 04С, опираясь локтями на колени, то сдвигает ноги, то раздвигает, таким образом движение, относящееся к сменам положения, оказывается объединяющим для нескольких каналов (руки и ноги), а 04R, также опираясь локтями на ноги, то поднимает колени, то опускает, что вызывает смену положения рук, головы и корпуса.

Описание мультиканальных кластеров позволяет также больше узнать об индивидуальных различиях между говорящими, таким образом мы получаем кинетические портреты участников записи: число каждого типа движений, их частота и сочетаемость в виде кластеров с движениями в других каналах представляют собой формально строгий и измеримый способ описания различий кинетического поведения. Так, у участника 23N больше половины движений ног, входящих в кластер с движениями рук, представляют собой движения-эхо (что вызвано опорой на ноги перед началом мануального жеста), а у 22С движения ног, входящие в кластеры с движениями рук, более чем в половине случаев — ритмические жесты (биты). Участник 04С чаще других, жестикулируя,

использует в одном кластере жесты рук и головы (5.2% цефалических жестов входят в кластер с движениями рук), а у O4R число таких кластеров, наоборот, меньше всех из изученных нами (0.2%).

Обсуждение и выводы

Жестикулируя, люди используют все привычные и доступные средства, и нередко жесты выполняются не только руками, но в одном жесте синхронно или последовательно задействованы руки и голова либо руки и ноги; по всей видимости, возможны и более сложные сочетания. Попытки описать их, опираясь на данные только одного канала, не отражают особенности реальной коммуникации.

При рассмотрении корпусных данных видно, что изучение одного канала отдельно от всех остальных часто приводит к неправильной трактовке, когда слишком слабые движения рук, но поддержанные более сложным жестом ног (например, при описании маршрута героев рассказа) будут проигнорированы или ошибочно отнесены к адапторам. Рассмотрение нетипичных жестов (локтей и ног) показывает, что поведение слушающего в разговоре устроено гораздо сложнее, чем мы привыкли считать. Традиционно считается, что жестикуляция – право говорящего, а слушающий не жестикулирует, он может только кивать, показывая свое внимание (McNeill, 1992). Однако нередки случаи, когда участник разговора, закончив свою реплику, 1–2 раза повторяет жесты ног или локтей той же самой формы, как когда они иллюстрировали его слова.

Еще одним применением предложенного подхода может служить описание индивидуальных особенностей жестикуляции – кинетические портреты.

Литература

Кибрик А. А., Федорова О. В. Эмпирическое исследование мультиканальной коммуникации: русские рассказы и разговоры о грушах // Психология. Журнал Высшей Школы экономики. 2018. Т. 15. № 2. С. 191 – 200.

Beattie G. Rethinking body language. How hand movements reveal hidden thoughts. London, New York: Routledge, 2016.

Chafe W. The pear stories: Cognitive, cultural, and linguistic aspects of narrative production. Norwood, New Jersey: Ablex, 1980.

Litvinenko A. O., Kibrik A. A., Fedorova O. V., Nikolaeva I. V. Annotating hand movements in multichannel discourse: Gestures, adaptors and manual postures // The Russian Journal of Cognitive Science. 2018. Vol. 5. No. 2. P. 4 – 17. <http://www.cogjournal.ru/5/2/pdf/LitvinenkoetalRJCS2018.pdf>

McNeill D. Hand and mind: What gestures reveal about thought. Chicago: University of Chicago Press, 1992.

Rehm M., André E. From annotated multimodal corpora to simulated human-like behaviors // Modeling Communication with Robots and Virtual Humans. 4930 Berlin, Heidelberg: Springer-Verlag, 2008. P. 1 – 17.

HANDS, LEGS AND HEAD: GOALS, DISCOVERIES AND CHALLENGES OF MULTICHANNEL ANNOTATION

Yu. V. Nikolaeva* (1), A. A. Evdokimova** (2)

* – julianikk@gmail.com, ** – arochka@gmail.com

1 – Lomonosov Moscow State University, Moscow; 2 – Institute of Linguistics, RAS, Moscow

Abstract. Considering communication in a multimodal perspective, there is a tendency toward isolation of channels: hand gestures are studied widely, and there are separate studies on head gestures; sometimes one of these channels is compared with verbal and/or prosodic data. However, there are few works about leg gestures, and comprehensive studies of different kinetic channels (hands head legs body facial expressions) are found only in relation to animated computer agents and in very limited contexts. Our work is aimed at developing a consistent annotation scheme for movements of different parts of the body within the kinetic modality, taking into account the specifics of each channel. This article discusses the phenomenon of multichannel clusters – a combination of movements in different channels to express a single meaning. We show that the generally accepted tendency to study the movements of only one channel leads to ignoring some significant gestures and to difficulties in describing natural communication.

Keywords: multichannel annotation, gesticulation, head movements, torso, leg and shoulder movements, multichannel clusters