

А. В. Варганов, И. В. Пасечник

ФЕНОМЕН СЕМАНТИЧЕСКОГО КОНФЛИКТА НА МОДЕЛИ ИСКУССТВЕННО СФОРМИРОВАННЫХ БИМОДАЛЬНЫХ ЗНАКОВ

В статье обсуждаются мозговые механизмы аудиовизуальной интеграции и понимания знака как прямого автоматического процесса (ранняя интеграция) и как процесса, требующего вторичной подстройки системы (поздняя интеграция) в ситуации семантического конфликта. Описывается формирование искусственных понятий с помощью специально разработанной системы бимодальных стимулов-знаков. Показано существование аудиовизуального семантического конфликта и выделены четыре стратегии его разрешения. Для каждой стратегии построены математические модели, описываемые линейными уравнениями, различающимися весами вкладов аудиальной и визуальной модальностей.

Ключевые слова: аудиовизуальная интеграция, формирование понятий, бимодальный знак, семантический конфликт.

We discussed the brain mechanisms of audiovisual integration and semantic comprehension as a direct process of comprehension (early integration) and process with second tuning (late integration) necessary in conflict situations. Existing of such audiovisual semantic conflict and four strategies to solve it were obtained in psychophysical experiment. Mathematic models were formulated according to that data. All models were described by linear equations, which were noted for the different weights of contributions of audio and visual modalities.

Key words: audiovisual integration, forming of concepts, bimodal sign, semantic conflict.

Проблемный контекст исследования

Мозговые механизмы функционирования знаковой (понятийной) системы человека, понимания и порождения речи, а также ряда сопутствующих когнитивных процессов давно интересуют психологов и психофизиологов. При этом проблема мозговых механизмов семантики и речи постепенно выходит на первый план. Она получила особое значение и в предложенной Е.Н. Соколовым теории «кодирования номером

Варганов Александр Валентинович — канд. психол. наук, ст. науч. сотр. кафедры психофизиологии ф-та психологии МГУ. *E-mail:* a_v_vartanov@mail.ru

Пасечник Инна Викторовна — науч. сотр. Института когнитивных исследований РНЦ «Курчатовский институт». *E-mail:* Passechnik@yandex.ru

Работа выполнена при финансовой поддержке РФФИ (грант № 06-06-80335).

канала». Задача обосновать общий, универсальный характер принципов кодирования, описанных первоначально на материале восприятия (Измайлов, Соколов, 1984; Соколов, Вайткявичус, 1989; Фомин и др., 1979), распространить их на все высшие психические функции человека была поставлена давно. В связи с этим в 1980-х гг. были начаты исследования семантики (первоначально на примере слов-названий цветов: Вартанов, Соколов, 1994, 1995; Соколов, Вартанов, 1987) и памяти (Вартанов и др., 1995). В настоящее время активно изучаются различные классы феноменов, связанных с интеграцией разномодальной информации (влияние контекста на понимание речевых высказываний, зависимость ошибок припоминания от условий и контекста запоминания).

Многие работы в этой области проведены на материале широко известного феномена (иллюзии) МакГурка—МакДональда (McGurk, MacDonald, 1976), обнаруженного в следующей экспериментальной ситуации. Испытуемым предъявлялся комбинированный стимул, состоящий из видеозаписи произнесения человеком слога «га» и аудиозаписи произнесения тем же человеком слога «ба». В итоге испытуемые слышали слог «да», которого реально не было ни в слуховой, ни в зрительной стимуляции¹. Такое влияние зрительной информации на слуховое восприятие было названо *аудиовизуальной интеграцией* (АВИ). Механизм этого влияния до конца не изучен. Остаются вопросы. Можно ли предсказать, в каком направлении трансформируется воспринимаемый звук под воздействием зрительной стимуляции? Может ли слуховая информация повлиять на зрительное восприятие? На каком уровне переработки информации происходит АВИ и какими физиологическими процессами она обеспечивается?

Один из методов изучения мозговых механизмов АВИ — сопоставление вызванных потенциалов (ВП), зарегистрированных на зрительные и слуховые стимулы, предъявляемые отдельно и вместе (бимодально). В качестве стимулов в исследованиях использовались семантические единицы разного уровня — фонемы, слоги, слова. Ставился вопрос: является ли АВИ результатом простого смешения стимулов разной модальности на сенсорном уровне или существует самостоятельный процесс интегрирования изначально раздельных слуховых и зрительных ощущений? В ряде экспериментальных работ (Besle et al., 2004; Kelly et al., 2004; Klucharev et al., 2003; Pourtois et al., 2000) было показано существование отдельного физиологического процесса, обеспечивающего АВИ, и вместе с тем высказано предположение о наличии так называемой ранней АВИ, происходящей на этапе восприятия. Эти результаты подтверждаются в исследовании физиологических механизмов эффекта МакГурка—МакДональда, проведенном в рамках «одд-болл парадиг-

¹ С иллюзией МакГурка—МакДональда можно ознакомиться в Интернете: <http://www.psy.msu.ru/illusion/others.html>

мы» (Colin et al., 2002): обнаружена негативность рассогласования на латенции 100—200 мс как для случая аудиовизуального восприятия (предполагающего процесс АВИ), так и отдельно для аудиального восприятия. Имеются аргументы и в пользу поздней интеграции (Kelly et al., 2004), возможно, связанной со специфической языковой обработкой информации (Klucharev et al., 2003).

Вопрос о времени разветвления процесса возникает и в связи со спецификой восприятия значения слова в разных контекстах. При этом дихотомия «ранняя/поздняя», отражающая *время* обработки речевой информации, соотносится с дихотомией «непосредственная (невербальная)/речевая (семантическая)», отражающей *уровень* ее обработки. Так, считается, что семантический анализ слова осуществляется позже (говорят о «когнитивном» компоненте N400), чем воспринимается его сенсорная оболочка. Обнаружено, например, что различие мозговых процессов, связанных с пониманием абстрактных и конкретных слов, проявляется в динамике именно поздних компонентов ВП — на латенции порядка 400 мс после подачи стимула (Friederici, 2004; Nittono et al., 2002). Однако имеются аргументы в пользу существования системы быстрого семантического анализа. Различия в мозговой активности, отражающие обработку слов и их семантический анализ, обнаруживаются на латенции порядка 100 мс (Skrandies, 1998; Worden, Foxe, 2003). В наших электрофизиологических исследованиях (Вартанов, Пасечник, 2005, 2006) также описаны мозговые механизмы раннего этапа понимания человеком значения слова-омонима. Показано, что амплитуда раннего компонента ВП P200 изменяется при изменении сложности семантического анализа.

Таким образом, можно признать, что существуют как ранние, так и поздние этапы АВИ. Ранний этап заключается во взаимном влиянии специфических сенсорных зон (Besle et al., 2004), тогда как поздний протекает в гетеромодальной коре. В ряде исследований, в том числе выполненных с помощью фМРТ, в качестве такой гетеромодальной коры указывается верхняя височная извилина (Callan et al., 2003; Jones, Callan, 2003; Macaluso et al., 2004; Mottonen et al., 2002; Sekiyama et al., 2003). Было обнаружено увеличение активности этой извилины в ситуации восприятия зашумленного звука при бимодальной стимуляции (Callan et al., 2003; Sekiyama et al., 2003). Есть несколько моделей поздней интеграции. Автор одной из них (Massaro, 1998) полагает, что интеграция двух модальностей происходит на основе сличения отдельных частей бимодального стимула с некоторым внутренним образом. Существует и моторная модель (Gentilucci, Cattaneo, 2005), исходящая из того, что для восприятия элементов речи (особенно в ситуации, когда она зашумлена) необходимо скрытое повторение речевых движений; при смешении этих скрытых артикуляторных движений происходит интеграция. Здесь предполагается, что речь должна быть достаточно развита.

Однако имеются данные, подтверждающие существование внеречевой АВИ, проявляющейся на сенсорном уровне с простейшими не-семантическими стимулами. Так, было показано, что если без всякого предварительного обучения сопровождать одну вспышку света двумя или тремя щелчками, то испытуемый переоценивает количество вспышек (Andersen et al., 2004; Shams et al., 2000, 2002). Кроме того, в ряде работ иллюзия МакГурка—МакДональда была обнаружена у 6-месячных младенцев, еще не владеющих речью (Burnham, Dodd, 2004; Rosenblum et al., 1997). Это указывает на существование врожденных механизмов более низкого (досемантического) уровня межмодального взаимодействия и АВИ. Изучение процессов аудиовизуальной интеграции в речи, их искажения, а также специфики их протекания в патологии (de Gelder et al., 2003; Williams, Massaro, 2004) показало, что у больных с поражениями стриарной коры сознательное восприятие эмоционального выражения лица или эмоциональной картинки вовсе не является обязательным условием для возникновения АВИ с эмоционально-выразительным голосом (Gelder et al., 2002). Этот вывод подтвержден на физиологическом уровне — уменьшением амплитуды слухового ВП. Исследование возрастных особенностей эффекта МакГурка—МакДональда (Cienkowski, Carney, 2002) выявило следующее: в ситуации когнитивного конфликта, в процессе принятия решения молодые люди чаще опираются на звуковую стимуляцию; с возрастом опора смещается в сторону визуальной стимуляции. В ситуации конфликта при бимодальном предъявлении речи человек делает выбор в пользу наименее неопределенного стимула той модальности, которая преимущественно характерна для данной возрастной категории.

Однако во всех вышеописанных экспериментах обычно использовались методики дихотомического ответа, которые не позволяли получить промежуточные варианты и количественно описать феномен АВИ. Вследствие этого было невозможно предсказать содержание возникающей иллюзии, т.е. объяснить механизм взаимовлияния модальностей при восприятии конфликтного стимула. Таким образом, хотя данных по аудиовизуальной интеграции, включая семантический уровень, а также по эффекту МакГурка—МакДональда в целом достаточно много, общие механизмы АВИ этого феномена остаются невыясненными.

Цель настоящей работы заключалась в исследовании механизмов функционирования понятий, предполагающих наличие связи между означаемым и означающим, т.е. взаимодействия соответствующих локальных анализаторов. Применение специальной системы бимодальных знаков позволяет использовать феномены АВИ и организовать управляемый конфликт. Для этого была разработана специальная методика формирования понятия и градуальной оценки его значения, а также количественного описания способов решения конфликта.

Методика

Процедура. Эксперимент заключался в формировании искусственных понятий на основе простой ассоциации двух наборов стимулов — означающих и означаемых. В качестве означающего использовались специально сформированные бимодальные стимулы, означаемым служили цветовые стимулы. Формирование проводилось посредством чередования обучающей и проверочной серий. *Обучающая серия:* испытуемому в случайном порядке предъявляются по 3 раза все ассоциации означающих и означаемых стимулов. *Проверочная серия:* испытуемому предъявляется стимул-означающее, к которому надо выбрать из ряда стимулов-означаемых правильную ассоциацию. Обратная связь о правильности или неправильности ответов испытуемому не давалась. Серии повторялись несколько раз до тех пор, пока испытуемый не достигал полной уверенности в правильном усвоении материала. Только после усвоения испытуемому давались 3 проверочные серии, дополнительно включающие 2 конфликтных стимула (принцип составления конфликтных стимулов будет описан ниже). Фиксировались ответы испытуемых во всех проверочных сериях — с конфликтными стимулами и без них.

Стимуляция. Бимодальные стимулы образовывались сочетанием двух градуально варьируемых качеств: 1) визуальная часть задавалась размером серого квадрата (9 градаций со стороны от 3 до 14 см), экспонируемого в течение 500 мс в центре экрана монитора компьютера; 2) аудиальная часть задавалась звуковыми интервалами (аккордами), образованными из двух чистых (синусоидальные колебания) тонов, синтезированных на компьютере: постоянный нижний тон — 180 Гц и изменяющийся по высоте второй тон (9 градаций от 200 до 300 Гц, длительность 200 мс).

В эксперименте использовались два *конфликтных стимула* (КС). В одном из них квадрат из стимула 3 сочетался со звуковым интервалом из стимула 5 (условное обозначение — КС[3+5]), во втором квадрат из стимула 5 сочетался со звуковым интервалом из стимула 3 (условное обозначение — КС[5+3]). Такие КС были основаны на возможности человека к АВИ, однако проявляется этот конфликт и в качестве семантического, т.к. означающему должно соответствовать какое-либо означаемое, а каждая из составных частей стимула в обучении соответствовала разным означаемым. Таким образом, в данном случае можно говорить о семантическом аудиовизуальном конфликте.

Означаемое задавалось в виде оттенков цветов равной яркости и насыщенности в диапазоне от красного к зеленому через желтый. Для проверочной серии было подобрано 50 оттенков, образующих градуально изменяющуюся цветовую систему, равномерно заполняющую весь экран монитора компьютера. Цветовая таблица состояла из вертикально расположенных цветных полосок, еле заметно изменяющихся слева на-

право от красного до зеленого. При этом только девять оттенков цветов являлись «правильными» означаемыми, т.е. соответствовали определенным бимодальным знакам, а другие были нужны для усложнения задачи испытуемому и определения величины ошибки означения. В проверочной серии испытуемый, перемещая курсор с помощью компьютерной мыши, должен был найти на палитре оттенок, означаемый предъявленным бимодальным стимулом. Каждому цвету на данной палитре приписывалось цифровое значение (в условных единицах перемещения курсора мыши), так что «самый красный» цвет (левый край монитора) соответствовал минимальному значению, а «самый зеленый» (правый край монитора) — максимальному.

Градации всех трех качеств (два качества означающего и качество означаемого) выбирались так, чтобы сохранить субъективное равенство шагов. Данные параметры стимулов определялись на основе предварительного психофизического эксперимента с попарным шкалированием различий большого набора стимулов: была показана одномерность обеих систем — и квадратов, и аккордов, отобранные стимулы образовывали соответствующие линейные функции. Параметры цветовых стимулов тоже выбирались (на основе проведенных ранее экспериментов с попарным шкалированием цветовых различий) так, чтобы возникла линейная зависимость воспринимаемого цветового качества от местоположения соответствующих цветовых образцов. В итоге все градации выбранных качеств оказались линейно связанными друг с другом и с цифровым значением местоположения проверочного материала (цветов-означающих) на экране компьютера.

Ассоциации составлялись так, чтобы выполнялась следующая закономерность: чем *больше* квадрат и сопровождающий его звуковой интервал, тем «более зеленому» стимулу соответствует данный бимодальный знак. В дополнительном эксперименте некоторым испытуемым предлагалось обратное соотношение величины квадрата с высотой звука — чем *больше* квадрат и *меньше* сопровождающий его звуковой интервал, тем «более зеленому» стимулу соответствует данный бимодальный знак.

Испытуемые. В эксперименте приняли участие 24 женщины и 7 мужчин в возрасте 20—65 лет. Еще двое мужчин и две женщины участвовали в дополнительном эксперименте.

Результаты

В каждой проверочной серии (с КС и без них) были получены оценки (в условных единицах, однозначно соответствующих оттенкам цвета заданной линейки спектра), являющиеся цифровыми эквивалентами значения знака. Для каждого испытуемого строились графики зависимости этих оценок от номера бимодального стимула (номер стимула определялся заданными величинами его составляющих). При сравнении

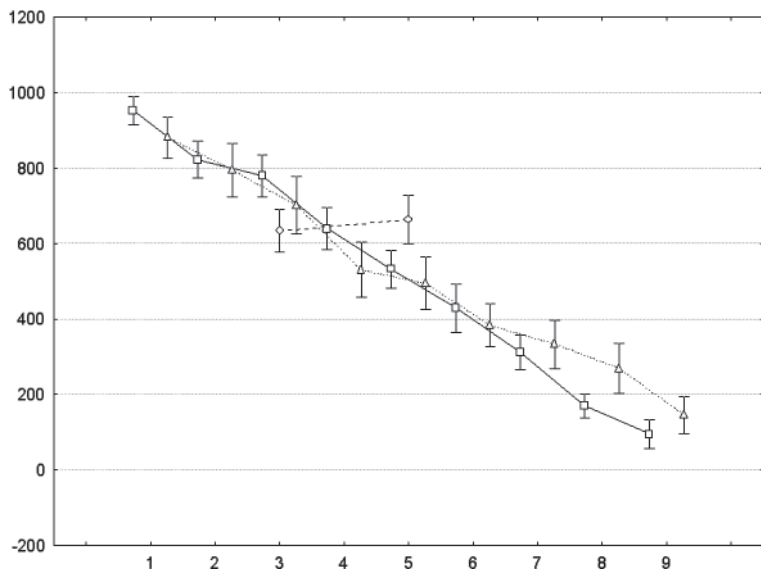


Рис. 1. Графики усредненных значений ответов испытуемых, reagировавших на конфликтные стимулы (КС) по первому типу (группа I, $n=7$). Здесь и на рис. 2, 3, 4: по абсциссе — номера бимодальных знаков, по ординате — цифровые эквивалентны выбираемых цветовых стимулов; условные обозначения: КС — ромб, штриховая линия; обычные стимулы в серии с КС — квадрат, сплошная линия; обычные стимулы в серии без КС — треугольник, пунктирная линия; скобочками показаны 95%-ные доверительные интервалы

распределений данных в сериях с КС и без них, т.е. по типу реакции на аудиовизуальный конфликт, все испытуемые были разделены на четыре группы. Следует отметить, что групповые графики для обычных (бесконфликтных) стимулов хорошо описываются линейной функцией (как и ожидалось). При этом для каждой группы испытуемых удалось построить линейную модель аудиовизуального конфликта, т.е. для всех бимодальных стимулов (и обычных, и конфликтных) подобрать единые коэффициенты аудиальной (A_{BC}) и визуальной (B_{BC}) составляющих, что позволяет предсказать результат АВИ для каждого испытуемого в зависимости от значений этих составляющих.

Проанализируем графики усредненных (групповых) значений ответов испытуемых, reagировавших на КС по тому или иному типу.

В **группе I** (7 испытуемых, рис. 1) наблюдается приближение значений обоих КС к значению стимула 4 (среднего между стимулами 3 и 5). Иными словами, КС[3+5] оценивается испытуемыми как «более зеленый», чем стимул 3 (различие достоверно по Т-критерию Стьюдента на уровне $p < 0.001$), а КС[5+3] как «более красный», чем стимул 5 (различие достоверно по Т-критерию Стьюдента на уровне $p < 0.002$). Также

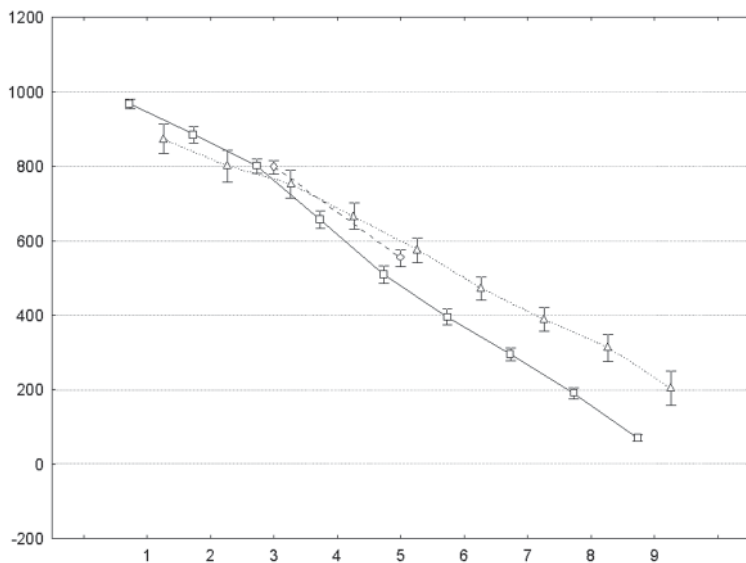


Рис. 2. Графики усредненных значений ответов испытуемых, reagировавших на конфликтные стимулы по второму типу (группа II, n=16)

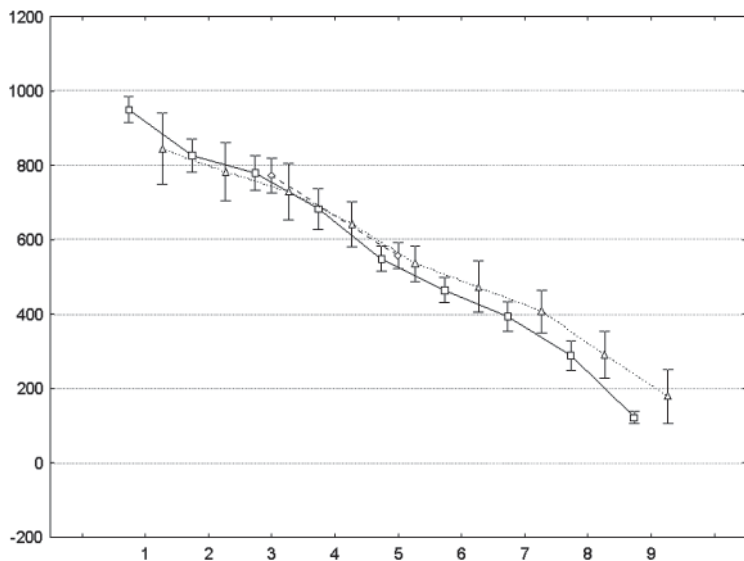


Рис. 3. Графики усредненных значений ответов испытуемых, reagировавших на конфликтные стимулы по третьему типу (группа III, n=4)

можно отметить, что графики оценок стимулов в сериях с КС и без них практически совпадают, значения расходятся только на стимулах 7, 8 и 9. Для этой группы коэффициент корреляции фактических и рассчитанных оценок в соответствии с линейной моделью равен 0.808, а модель характеризуется тем, что коэффициент влияния $A_{BC} = 61.723 \pm 9.30$ ($p < 0.0010$), а $B_{BC} = 47.185 \pm 9.30$ ($p < 0.001$).

Для **группы II** (16 человек, рис. 2) характерно совпадение значений стимула 3 и КС[3+5], но КС[5+3] оценивается как «более красный», чем стимул 5. Графики значений обычных стимулов в сериях с КС и без них совпадают на первых трех стимулах, но в дальнейшем расходятся — по остальным стимулам наблюдается все более сильная переоценка в сторону «более зеленого» цвета в серии с КС в сравнении с серией без КС. Для данной группы коэффициент корреляции фактических и рассчитанных оценок равен 0.906. В этой модели коэффициент влияния $A_{BC} = -2.576 \pm 3.576$ ($p = 0.471$, т.е. отличие от 0 данного коэффициента недостоверно), а $B_{BC} = 119.262 \pm 3.575$ ($p < 0.001$).

4 испытуемых, составивших **группу III** (рис. 3), игнорировали A_{BC} , оценивая значения обоих КС только по соответствующим B_{BC} , т.е. КС[3+5] воспринимали как стимул 3, а КС[5+3] — как стимул 5. Для этой группы характерно совпадение значений обычных стимулов в сериях с КС и без КС. Коэффициент корреляции фактических и рассчитанных оценок равен 0.911. В этой модели коэффициент влияния

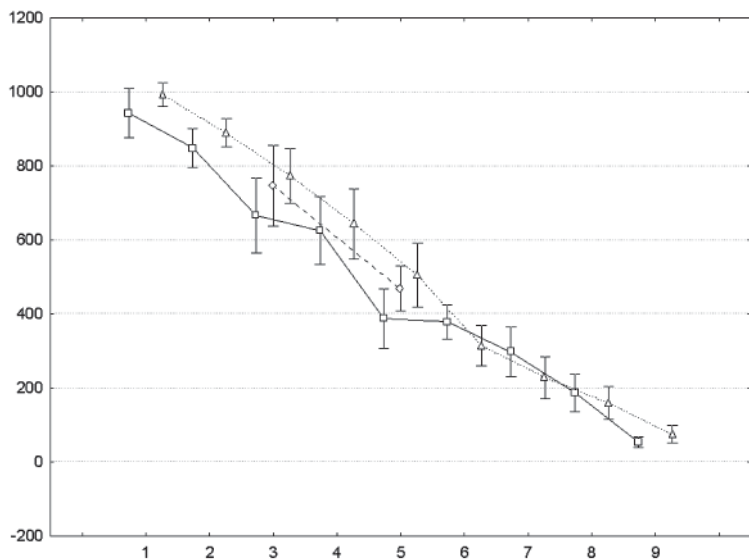


Рис. 4. Графики усредненных значений ответов испытуемых, reagировавших на конфликтные стимулы по четвертому типу (группа IV, n=4)

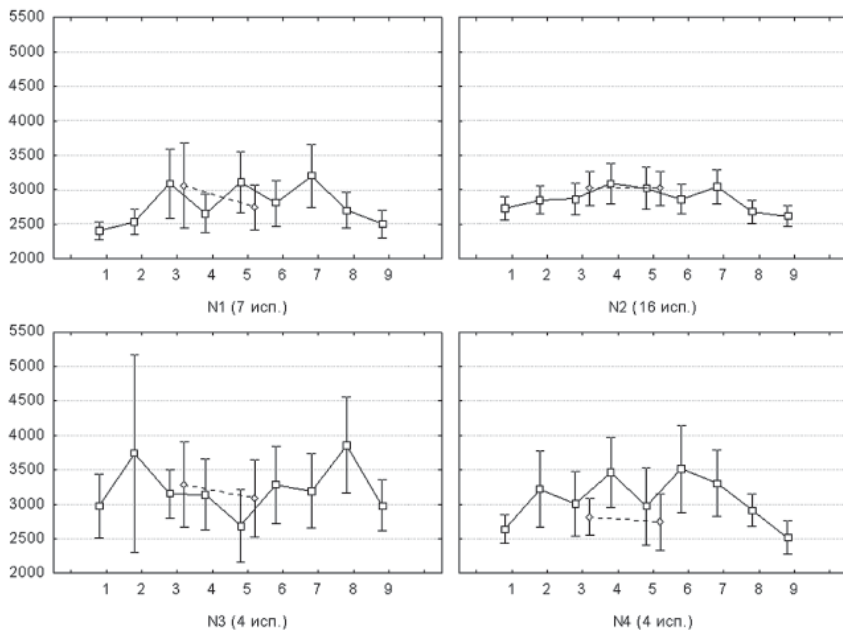


Рис. 5. Графики времени реакции (BP) на обычные и конфликтные стимулы в четырех группах испытуемых. По абсциссе — номера стимулов, по ординате — BP в мс. Условные обозначения: обычные стимулы — квадрат, сплошная линия; конфликтные стимулы — ромб, штриховая линия; скобками показаны 95%-ные доверительные интервалы

$A_{BC} = -3.650 \pm 7.083$ ($p=0.607$, т.е. отличие от 0 данного коэффициента недостоверно), а $B_{BC} = 102.367 \pm 7.083$ ($p < 0.001$).

Наконец, в **группе IV**, состоящей из 4 человек (рис. 4), отмечается недооценка значений (восприятие их как «более красных») обоих КС по сравнению с соответствующими обычными стимулами. Графики значений обычных стимулов для серий с КС и без КС значительно расходятся по первым пяти стимулам, причем значения стимулов в серии с КС сдвинуты в сторону зеленого. Для данной группы коэффициент корреляции фактических оценок с рассчитанными по модели равен 0.918. В этой модели коэффициент влияния $A_{BC} = -2.455 \pm 6.864$ ($p=0.001$), а $B_{BC} = 134.827 \pm 6.864$ ($p < 0.001$).

В каждой из групп было оценено время реакции на конфликтные и обычные стимулы. На рис. 5 видно, что ни в одной группе нет существенных различий в скорости оценки тех и других стимулов, выявляется только общая закономерность увеличения времени реакции на стимулы срединного (наиболее трудного) диапазона.

Испытуемые, участвовавшие в дополнительном эксперименте (когда направление изменения аудиального качества, соответствующее

изменению визуального качества, было инвертировано), показали такие же результаты, что и испытуемые основной группы. Это говорит об универсальности полученной закономерности.

Обсуждение

Попытки градуально дозировать аудиальную и визуальную составляющие в эффекте МакГурка—МакДональда предпринимались и раньше. В частности Д. Массаро (Massaro, 1998) при помощи специальной программы создал континуум звуков и мимики при переходе от слога «га» к слогу «ба». Однако реакция испытуемого оставалась бинарной: он мог ответить, что слышал либо «ба», либо «да», хотя и с разной вероятностью в зависимости от вклада соответствующих модальностей.

В нашем эксперименте использование системы искусственно сформированных знаков-стимулов, образованных сочетанием двух градуально варьируемых качеств, дало возможность перейти от анализа бинарной вероятности разрешения аудиовизуального конфликта, аналогичного эффекту МакГурка—МакДональда, к анализу характера распределения количественных оценок, реализующих его разрешение.

Искусственность системы знаков исключала влияние предыдущего опыта и личностных предпочтений испытуемых: все эти знаки воспринимались ими одинаково, даже конфликтные стимулы очень редко (лишь в конце эксперимента) и отдельными испытуемыми осознавались как «более искусственные» и более трудные для оценки.

Результаты сопоставления индивидуальных данных, выделение четырех групп испытуемых позволили выдвинуть предположение о наличии разных типов усвоения материала и стратегий разрешения аудиовизуального конфликта. По-видимому, различие стратегий объясняется индивидуальными особенностями испытуемых, которые удалось выявить благодаря выбранной процедуре обучения и примененной экспериментальной схеме. В данном эксперименте использовалось обучение без обратной связи: испытуемому не сообщалось, правильно или нет (или насколько точно) он указал цвет, соответствующий бимодальному стимулу. При этом трудности в определении КС, по-видимому, могли переживаться испытуемым как показатель недостаточной обученности, а это стимулировало его к продолжению поиска более подходящей закономерности в соотношении означающего и означаемого с учетом не только обычных стимулов, но и КС. Некоторых испытуемых наличие КС могло привести к тому, что обучение в этой ситуации становилось фактически обучением с неправильной обратной связью.

Как следует из самоотчетов, испытуемые группы I, по сравнению с испытуемыми других групп, при усвоении бимодальных знаков в гораздо большей степени учитывали аудиальную информацию. В некоторых случаях они замечали КС, но не пересматривали всю сформированную на обычных стимулах знаковую систему. Подтверждением такой

субъективной стратегии стала построенная математическая модель, согласно которой линейные коэффициенты вкладов A_{BC} и B_{BC} достаточно близки, что означает практически равный вклад информации обеих модальностей в определение значения стимула. По графику (см. рис. 1) можно видеть, что определение значений КС происходит независимо от значений бесконфликтных стимулов, включенных в систему; значение последних не меняется в зависимости от того, предъявляются наравне с ними КС или нет.

В группах II и IV КС, как правило, не распознаются (и не осознаются), но учитываются в ходе обучения, продолжающегося даже в проверочных сериях. При этом происходит подстройка всей системы знаков в целом, отражающаяся в расхождении значений обычных стимулов в сериях с КС и без КС, т.е. в преувеличении значений стимулов в серии с КС в сторону зеленого цвета. Из самоотчетов этих испытуемых следует, что они в большей степени опирались на визуальную информацию, но звук слышали и пытались на него ориентироваться. В этих группах рассчитанные коэффициенты для математической модели отражают значительную (подавляющую) роль визуальной информации, но обнаруживают различия по коэффициентам вклада аудиальной информации для каждой из групп. В группе II коэффициент вклада аудиальной информации практически сводится к нулю, тогда как в группе IV он значимо отличается от нуля, но является отрицательным. Можно сказать, что в группе IV аудиальная информация «тормозила» или закономерно преуменьшала соответствующую визуальную информацию.

Наконец, в самоотчетах испытуемых группы III достаточно ясно говорится об изначальном и полном игнорировании аудиальной информации, как на стадии обучения, так и на стадии проверки, вследствие чего у этих испытуемых и не было необходимости продолжать обучение и менять общую настройку знаковой системы, сложившейся у них в обучающих и первых проверочных сериях. Этот вывод следует из факта совпадения графиков обычных стимулов в сериях с КС и без КС. Испытуемые этой группы были вполне успешны в обучении, так как в обычных стимулах аудиальная информация полностью дублировала информацию о значении стимула, которая содержалась в визуальной части стимула и была, по сути, избыточной. Однако ориентировка вопреки инструкции на неполную информацию понижала их чувствительность и способность к обучению в более сложных условиях.

Результаты эксперимента позволяют говорить о влиянии аудиальной информации на визуальную, так как в связи с особенностью стимульного материала ведущей модальностью все-таки оказывалось зрение, а слух был лишь средством «тонкой настройки» знаковой системы. Выделенные нами группы в основном различаются коэффициентами вклада каждой модальности в общее восприятие. Только

испытываемые группы I учитывали обе части бимодального стимула в приблизительно равной мере, что позволяет точно предсказать качество воспринимаемого КС, рассчитываемое как среднее арифметическое A_{BC} и B_{BC} . Конфликтный бимодальный стимул они воспринимали как новый, не совпадающий ни с одной из его составляющих. Этот случай полностью аналогичен эффекту МакГурка—МакДональда, когда испытуемые слышали «да», а не «ба» или «га». Если бы от них потребовали бинарного ответа, то его вероятность определялась бы соотношением аудиальной и визуальной составляющих. Однако в отношении данного эффекта следует говорить скорее о влиянии визуальной информации на аудиальную, чем наоборот, как в нашем примере, поскольку именно зрительное восприятие движений губ вызывало изменение слухового образа. Испытуемых групп II и IV наличие КС приводит к трансформации всей системы оценок, хотя и в этом случае можно точно рассчитать качество воспринимаемого КС как среднее арифметическое взвешенных с заданными весами A_{BC} и B_{BC} . В результате, под влиянием аудиальной информации, происходит переоценка B_{BC} (т.е. величины квадрата) по цветовой шкале в сторону зеленого. Возможность такой перестройки значений всех стимулов системы согласуется с данными о системном влиянии аудиальной информации на зрительное восприятие (Bertelson, Vroomen, de Gelder, 2003).

Итак, результаты, полученные с помощью предложенной экспериментальной схемы, позволили впервые количественно описать и предсказать специфику реагирования на аудиовизуальный семантический конфликт четырех групп испытуемых. Для каждой из этих групп рассчитывались собственные коэффициенты вкладов A_{BC} и B_{BC} , и было обнаружено существование универсальной линейной модели, подходящей для описания всех данных в целом.

Можно провести некоторую аналогию между вышеописанным экспериментом и данными, полученными в школе И.П. Павлова (Нарбутович, Подкопаев, 1936) по сочетанию зрительного и аудиального (первоначально индифферентных) стимулов при выработке условных рефлексов у собак. Было показано, что после выработки рефлекса на стимул одной модальности выработка аналогичного рефлекса на стимул второй модальности существенно ускорялась. По-видимому, формирование такой ассоциативной связи предполагает сходный механизм влияния анализатора одной модальности на анализатор другой модальности.

В соответствии с примененной экспериментальной схемой предполагалось, что в ходе обучения у испытуемых происходит формирование бимодального знака как целостности. Как следует из самоотчетов, практически все испытуемые не осознавали наличия конфликтных стимулов в предъявляемой системе бимодальных означающих. Остается вопрос, действительно ли в условиях проведенного обучения формировалась

целостность («гештальт») или разномодальные качества, составляющие означающее, независимо друг от друга ассоциировались со стимулом-означаемым (цветом)? Такие два варианта установления ассоциации подразумевают и различные нейронные механизмы обработки информации: в первом случае должно происходить взаимовлияние анализаторов (возможно, на стадии сенсорной обработки), во втором случае происходит параллельная обработка информации разных модальностей с последующим сличением на завершающей стадии результатов обработки информации каждым независимым каналом. Второй механизм АВИ описан в работе Д. Массаро (Massaro, 1998). При этом возможно независимое восприятие разномодальных составляющих стимула, а в случае конфликта возможно существование той или иной стратегии их «примирения». Таким образом, обсуждаемые результаты могут быть одинаково хорошо описаны как первым, так и вторым механизмом взаимодействия анализаторов. Совокупность полученных данных заставляет предположить наличие дополнительной настройки восприятия стимула путем повторного входа. Возможно, первичная семантическая обработка проходит параллельно с обработкой сенсорной информации о стимуле. Однако при получении информации о рассогласовании поступившего сигнала и имеющихся образов означающего (для групп II и IV) происходит повторная настройка всей системы (испытуемые групп I и III не ощущали необходимости каких-либо изменений всей системы). Это позволяет объяснить наличие как «ранней», так и «поздней» интеграции: функционирование уже сформированной системы вполне может проходить автоматически и очень быстро при непосредственном восприятии знака и «понимании» его значения, поскольку для этого не требуется обратной связи. Если же возникает необходимость корректировки системы, то она происходит на основе обратной связи и оценки эффективности функционирования системы, т.е. на поздних этапах интеграции. Это предположение хорошо согласуется не только с результатами нашего эксперимента, но и с литературными данными. Так, в одной из работ (Kelly, Kravitz, Hopkins, 2004) сопоставлялись мозговые процессы при восприятии речи на слух в сопровождении жестов и было показано, что при совпадении и несовпадении жеста и речи обнаруживаются различия как на позднем этапе интеграции (N400 с правой латерализацией), так и на ее раннем этапе билатерально в затылочных и фронтальных областях.

Выводы

1. Показано существование феномена семантического аудиовизуального конфликта на модели искусственно сформированной системы бимодальных знаков.
2. Выявлены различные типы приспособительных стратегий испытуемых в ситуации разрешения семантического конфликта.

3. Установлено, что процесс аудиовизуальной интеграции (вне зависимости от типа используемой стратегии) количественно описывается при помощи линейной модели с различными весами вкладов аудиальной и визуальной составляющих бимодального стимула.

СПИСОК ЛИТЕРАТУРЫ

Вартапов А.В., Манукян Н.К., Соколов Е.Н., Тсаконас К.Г. Сохранение цветового образа в кратковременной памяти // Журн. ВНД. 1995. Т. 45. № 6. С. 1085—1093.

Вартапов А.В., Пасечник И.В. Мозговые механизмы семантического анализа слов-омонимов // Журн. ВНД. 2005. Т. 55. Вып. 2. С. 193—197.

Вартапов А.В., Пасечник И.В. Психофизиологические особенности восприятия знака в условиях семантического конфликта // Мат-лы Междунар. науч.-практ. конф. «Развитие научного наследия Бориса Михайловича Теплова в отечественной и мировой науке (к 110-летию со дня рождения)». М., 2006. 76—80.

Вартапов А.В., Соколов Е.Н. Семантическое пространство цветовых названий: опыт межъязыкового исследования // Психол. журн. 1994. Т. 15. № 5. С. 56—66.

Вартапов А.В., Соколов Е.Н. Роль первой и второй сигнальных систем в соотношении семантического и перцептивного цветовых пространств // Журн. ВНД. 1995. Т. 45. № 2. С. 343—357.

Измайлов Ч.А., Соколов Е.Н. Цветовое зрение. М., 1984.

Нарбутович И.О., Подкопаев Н.А. Условный рефлекс как ассоциация // Труды физиологических лабораторий акад. И.П. Павлова. 1936. Т. VI. Вып. 2. С. 5—25.

Соколов Е.Н., Вайткявичус Г.Г. Нейроинтеллект: от нейрона к нейрокомпьютеру. М., 1989.

Соколов Е.Н., Вартапов А.В. К исследованию семантического пространства // Психол. журн. 1987. № 2. С. 58—65.

Фомин С.В., Соколов Е.Н., Вайткявичус Г.Г. Искусственные органы чувств. М., 1979.

Andersen T.S., Tiipana K., Sams M. Factors influencing audiovisual fission and fusion illusion // Cogn. Brain Res. 2004. Vol. 21. N 3. P. 301—308.

Bertelson P., Vroomen J., Gelder B., de. Visual recalibration of auditory speech identification: a McGurk aftereffect // Psychol. Sci. 2003. N 14. P. 592—597.

Besle J., Fort A., Delpuech C., Giard M.-H. Bimodal speech: early suppressive visual effects in human auditory cortex // Eur. J. Neurosci. 2004. Vol. 20. N 8. P. 2225—2234.

Burnham D., Dodd B. Auditory-visual speech integration by prelinguistic infants: perception of an emergent consonant in the McGurk effect // Dev. Psychobiol. 2004. Vol. 45. N 4. P. 204—220.

Callan D.E., Jones J.A., Munhall K. et al. Neural processes underlying perceptual enhancement by visual speech gestures // Neuroreport. 2003. Vol. 14. N 17. P. 2213—2218.

Cienkowski K.M., Carney A.E. Auditory-visual speech perception and aging // Ear and Hearing. 2002. Vol. 23. N 5. P. 439—449.

Colin C., Radeau M., Soquet A. et al. Mismatch negativity evoked by the McGurk—MacDonald effect: a phonetic representation within short-term memory // Clin. Neurophysiol. 2002. Vol. 113. N 4. P. 495—506.

Gelder B., de, Pourtois G., Weiskrantz L. Fear recognition in the voice is modulated by unconsciously recognized facial expression but not by unconsciously recognized affective pictures // PNAS. 2002. Vol. 99. N 6. 4121—4126.

Gelder B., de, Vroomen J., Annen L. et al. Audiovisual integration in schizophrenia // Schizophrenia Res. 2003. Vol. 59. N 2. P. 211—218.

Gentilucci M., Cattaneo L. Automatic audiovisual integration in speech perception // Exp. Brain Res. 2005. Vol. 167. N 1. 1—10.

- Friederici A.D.* Event-related brain potential studies in language // *Curr. Neurol. Neurosci. Rep.* 2004. Vol. 4. P. 466—470.
- Jones J.A., Callan D.E.* Brain activity during audiovisual speech perception: an fMRI study of the McGurk effect // *Neuroreport.* 2003. Vol. 14. N 8. P. 1129—1133.
- Kelly S.D., Kravitz C., Hopkins M.* Neural correlates of bimodal speech and gesture comprehension // *Brain Lang.* 2004. Vol. 89. N 1. 253—260.
- Klucharev V., Mottonen R., Sams M.* Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception // *Cogn. Brain Res.* 2003. Vol. 18. N 1. 65—75.
- Macaluso E., George N., Dolan R. et al.* Spatial and temporal factors during processing of audiovisual speech: a PET study // *Neuroimage.* 2004. Vol. 21. N 2. P. 725—732.
- Massaro D.W.* *Perceiving talking faces: from speech perception to a behavioral principle.* Cambridge, MA, 1998.
- McGurk H., MacDonald J.* Hearing lips and seeing voices // *Nature.* 1976. Vol. 264. P. 746—748.
- Mottonen R., Krause C.M., Tiippana K., Sams M.* Processing of changes in visual speech in the human auditory cortex // *Cogn. Brain Res.* 2002. Vol. 13. N 3. P. 417—425.
- Nitto H., Suehiro M., Hori T.* Word imageability and N400 in an incidental memory paradigm // *Int. J. Psychophys.* 2002. Vol. 44. P. 219—229.
- Pourtois G., de Gelder B., Vroomen J. et al.* The time-course of intermodal binding between seeing and hearing affective information // *Neuroreport.* 2000. Vol. 11. N 6. P. 1329—1333.
- Rosenblum L.D., Schmuckler M.A., Johnson J.A.* The McGurk effect in infants // *Percept Psychophys.* 1997. Vol. 59. N 3. P. 347—357.
- Sekiyama K., Kanno I., Miura S., Sugita Y.* Auditory-visual speech perception examined by fMRI and PET // *Neurosci. Res.* 2003. Vol. 47. N 3. P. 277—287.
- Shams L., Kamitani Y., Shimojo S.* Illusions: What you see is what you hear // *Nature.* 2000. Vol. 408. P. 788—789.
- Shams L., Yukiyasu K., Shimojo S.* Visual illusion induced by sound // *Cogn. Brain Res.* 2002. Vol. 14. P. 147—152.
- Skrandies W.* Evoked potential correlates of semantic meaning — a brain mapping study // *Cogn. Brain Res.* 1998. Vol. 6. P. 173—183.
- Williams J.H., Massaro D.W.* Visual-auditory integration during speech imitation in autism // *Res. Dev. Disabil.* 2004. Vol. 25. N 6. P. 559—575.
- Worden M.S., Foxe J.J.* The dynamics of the spread of selective visual attention // *PNAS.* 2003. Vol. 100. P. 11933—11935.