

УДК 81'344.4+81'221.24

П. А. Скребин, У. Е. Кочеткова

г. Санкт-Петербург, Россия, Санкт-Петербургский государственный университет

ОСОБЕННОСТИ СИНХРОНИЗАЦИИ ЕДИНИЦ СУПРАСЕГМЕНТНОГО УРОВНЯ С НЕВЕРБАЛЬНЫМИ ЯВЛЕНИЯМИ В НЕЙТРАЛЬНОЙ И ЭМОЦИОНАЛЬНО-ОКРАШЕННОЙ РЕЧИ

Целью настоящей работы является анализ синхронизации мимики и жестов с границами супraseгментных единиц и информационным фокусом высказывания. Исследование проводилось на материале мультимедийного корпуса иронической речи, созданного на кафедре фонетики и методики преподавания СПбГУ. Результаты анализа показали различия в синхронизации интонационных и невербальных средств в зависимости от типа речи и наличия эмоциональной окраски.

Ключевые слова: акустический анализ; информационный фокус; интонационный центр; просодические характеристики; мелодический контур; мимика и жестикуляция; синхронизация.

P. A. Skrelin, U. Y. Kochetkova

St. Petersburg, Russia, St. Petersburg State University

FEATURES OF SYNCHRONIZATION OF PROSODIC UNITS WITH NONVERBAL PHENOMENA IN NEUTRAL AND EMOTIONAL SPEECH

The purpose of this work is to analyze the synchronization of facial expressions and gestures with the boundaries of prosodic units and the information focus of the utterance. The research was conducted on the material of the multimedia corpus of ironic speech created at the

Department of Phonetics and Teaching Methods of St. Petersburg State University. The results of the analysis showed differences in the synchronization of intonation and nonverbal means depending on the type of speech and the presence of emotional colouring.

Key words: acoustic analysis; information focus; nucleus; prosodic characteristics; melodic contour; facial expressions and gestures; synchronisation.

Известно, что невербальный канал передачи информации задействуется говорящим не менее активно, чем вербальный. Взаимодействию этих каналов посвящено большое количество исследований. В настоящее время результаты таких исследований становятся особенно важны в связи с возросшим количеством аудиовизуальных помощников, интерфейсов и антропоморфных роботов, поскольку разработчики вынуждены моделировать не только речевое поведение человека, но и сопутствующие ему паралингвистические характеристики. Особенно важно получить в этом случае правильное выражение различных эмоций и добиться естественности. Кроме того, интерес представляют и сочетания определенных акустических ключей с паралингвистическими, которые могут улучшить качество автоматического распознавания речи.

В ряде исследований, проведенных на материале спонтанной и подготовленной речи в различных языках, было показано, что синхронизация жестов, в том числе мимических, с информационным фокусом высказывания, который на просодическом уровне соотносится с его интонационным центром, больше характерна для подготовленной речи, нежели для спонтанной [1, 2, 3]. Кроме того, достаточно часто в подготовленной речи наблюдается параллелизм между направлением движения основного тона и направлением жеста [4, 5, 6, 7]. Некоторые работы говорят и о совпадении ударной части жеста с акустическим пиком высказывания [1].

Существуют работы, посвященные изучению жестов, характерных для эмоционально-окрашенных высказываний, в частности, при выражении иронического значения. В таких высказываниях среди жестикуляторов отмечают многократные (повторяющиеся) движение бровей [8] и/или головы [9]. В работе С. Гонсалес-Фуэнте и соавторов были получены данные о важной роли паралингвистических маркеров иронии при отсутствии контекста [10].

Несмотря на интерес исследователей к данной теме, работы по изучению синхронизации паралингвистических и просодических средств при выражении иронии в русском языке до сих пор отсутствовали. Первым таким исследованием стал анализ иронической актерской речи на материале российских фильмов и сериалов [11].

Целью же настоящей работы является установление характера взаимодействия жестов и мимики с просодическими характеристиками иронических высказываний в квази-спонтанной лабораторной речи, а также их синхронизация с границами синтагмы (в понимании Л. В. Щербы) и интонационного центра высказывания.

В качестве материала нами были использованы наборы омонимичных ироничных и неироничных целевых фрагментов и окружающих их контекстов из мультимедийного корпуса иронической русской речи, разработанного на кафедре фонетики СПбГУ в рамках проекта «Акустические характеристики иронии при реализации функциональных интонационных моделей», поддержанного грантом Российского фонда фундаментальных исследований № 20-012-00552.

Аудиозапись проводилась в звукоизолированной кабине одновременно с видеозаписью с частотой 100 кадров в секунду при наличии согласия со стороны диктора. Дикторами выступили носители русского языка. Запись предполагала чтение напечатанных на бумаге текстов и наборов фраз. С этим были связаны некоторые ограничения в использовании дикторами жестикуляторов, что привело к крайне редкому использованию движений руками и глазами, которые были бы зафиксированы на видеокамеру, стоящую напротив читающего. Дикторы могли ознакомиться с текстом перед прочтением, однако тренировка не предполагалась.

Для проведения настоящего исследования была использована часть мультимедийного корпуса: набор из 62 мини-текстов, состоящих из 2–4 фраз каждый, предполагающих ироническое и неироническое прочтение, в произнесении 12 дикторов (7 мужчин и 5 женщин). Таким образом были получены 744 стимула.

Акустический анализ производился с помощью программного обеспечения Praat, он позволил установить границы синтагм, ударных слогов и ударного гласного внутри этих слогов. Метки частоты основного тона расставлялись в программе Wave Assistant. Для анализа жестов и мимики использовалась программа Elan. Рассматривались следующие жестикуляторы: голова, брови, глаза, плечи, руки, губы. Кроме того, отмечалось такое паралингвистическое явление, как смех.

В ходе анализа жестикуляторов было обнаружено, что большинство как иронических, так и неиронических высказываний (соответственно 87 % и 86 %) характеризуются движением головы. Хотя многократные движения головой в целом уступают по частоте однократным в обоих типах высказываний, всё же можно заметить, что в иронических высказываниях такие многократные движения присутствуют в два раза чаще, чем в нейтральных высказываниях (соответственно, в 15 % и в 7 % случаев).

Мимические движения бровями чаще используются в иронических высказываниях (в 20 % случаев), чем в нейтральных (в 12 % случаев). В обоих типах высказываний возможны разнообразные однократные движения, в том числе подъем одной из бровей.

Движение глазами можно было проследить лишь у одного диктора из двенадцати. При этом подобные движения чаще присутствовали в нейтральной речи (в 15 % иронических и 20 % неиронических высказываний).

Движения руками встречалось также только у одного диктора, причем, опять же, такие движения чаще присутствовали в иронической, нейтральной речи.

Как и ожидалось, всеми дикторами использовались смех и улыбка для передачи иронического значения, что отличало иронические высказывания от нейтральных.

При исследовании совпадения жестов и мимики с границами синтагм в иронической и нейтральной речи наблюдалось более частое совпадение в нейтральной речи (85 % высказываний), нежели в речи иронической (57 % высказываний). Это может быть связано с тем, что эмоциональное – в данном случае ироническое – значение часто распространяется и на контекст, окружающий целевое высказывание. С другой стороны, можно отметить и другое явление – смещение жеста, выражающего иронию, с целевого фрагмента на предшествующую или последующую паузу.

Анализ темпорального соответствия жестикуляции и информационного центра высказывания показал, что в большинстве иронических высказываний (70 %) наблюдалась синхронизация мимики и жестов с информационным фокусом и интонационным центром высказывания. Однако для нейронических высказываний также наблюдалось подобное совпадение (в 72 % высказываний). Это свидетельствует о том, что подобная синхронизация не может считаться отличительной чертой иронической речи.

Другая черта, общая для иронических и нейронических высказываний, – это синхронизация жеста с первым слогом высказывания, что несколько чаще наблюдалось в иронической речи (в 40 % случаев), нежели в нейронической (в 32 % случаев). Некоторые высказывания характеризовались наличием нескольких жестов, один из них совпадал с первым слогом, второй (один из следующих) – с информационным и интонационным центром.

Синхронизация жестов с первым ударным слогом (в том случае, если высказывание состояло из двух и более слов) оказалась более свойственной нейтральной, нейронической речи. Она наблюдалась в 72 % нейтральных высказываний и только в 45 % иронических высказываний.

В некоторых случаях наблюдалась синхронизация с постцентром высказывания – в иронических высказываниях чаще (в 18 % случаев), чем в нейронических (в 9 % случаев).

При сравнении направления движения ЧОТ и направления жестов оказалось, что совпадение в направлении движения на информационном (интонационном) центре гораздо чаще встречается в нейтральных высказываниях (в 67 % случаев), чем в иронических (в 25 % случаев). Эти данные отличаются от тех, что были получены на материале русской актерской речи, где во всех нейтральных и иронических высказываниях наблюдалось совпадение направления движения хотя бы одного из жестикуляторов с направлением движения тона на интонационном центре [11].

Подводя итоги, можно сказать о том, что характер взаимодействия мимики и жестикуляции во многом зависят от типа речи (подготовленная актерская речь и квази-спонтанная лабораторная речь) и наличия эмоциональной окраски, т.е. в нашем исследовании – наличия иронического значения. Однако ряд характеристик являются общими для иронических и нейтральных высказываний в обоих типах речи, например, синхронизация с информационным (интонационным) центром высказывания.

ЛИТЕРАТУРА

1. *Nobe S.* Representational gestures, cognitive rhythms, and acoustic aspects of speech: A network/threshold model of gesture production : Unpubl. diss. University of Chicago, 1996. 113 p.
2. *Nobe S.* Where the most spontaneous representational gestures actually occur with respect to speech? // *Language and Gesture* Cambridge University Press / McNeill, David (Ed.). 2000. P. 186–198.
3. *Wagner P., Malisz Z., Kopp S.* Gesture and Speech in Interaction: An Overview // *Speech Communication*. 2014. Vol. 57. P. 209–232.
4. *Bolinger D.* Intonation and gesture // *American Speech*. 1983. Vol. 58. P. 156–174.
5. Three-dimensional articulatory modeling of tongue, lips and face, based on MRI and video images / P. Badin, G. Bailly, L. Revéret, M. Baciú, C. Segebarth, C. Savariaux // *Journal of Phonetics*. 2002. Vol. 30, № 3. P. 533–553.
6. Visual prosody: facial movements accompanying speech / H. P. Graf, E. Cosatto, V. Strom, F. J. Huang // *Proceedings of AFGR*. 2002. P. 96–102.
7. Visual prosody and speech intelligibility / K. G. Munhall, J. A. Jones, D. E. Callan, T. Kuratate, E. Vatikiotis-Bateson // *Psychological Science*. 2004. Vol. 15, № 2. P. 133–137.
8. *Williams J. A., Burns E. L., Harmon E. A.* Insincere utterances and gaze: Eye contact during sarcastic statements // *Perceptual and Motor Skills*. 2009. Vol. 108. P. 565–572.
9. Multimodal markers of irony and sarcasm / S. Attardo, J. Eisterhold, J. Hay, I. Poggi // *Humor : International Journal of Humor Research*. 2003. Vol. 16, № 2. P. 243–260.
10. *González-Fuente S., Prieto P., Noveck I.* A fine-grained analysis of the acoustic cues involved in verbal irony recognition in French // *Proceedings of the International Conference on Speech Prosody*, 2016-January. Boston, 2016. P. 902–906.
11. Interplay of Visual and Acoustic Cues of Irony Perception: a Case Study of Actor's Speech / U. Kochetkova, V. Evdokimova, P. Skrelin, R. German, D. Novoselova // *Artificial Intelligence and Natural Language*. Cham, 2022. Vol. 1731. P. 82–94.