

Фонетические проявления коммуникативной адаптации в диалоге

© 2021

Татьяна Васильевна Качковская[@]

Анна Дмитриевна Мамушина

Санкт-Петербургский государственный университет, Санкт-Петербург, Россия;
kachkovskaia@phonetics.pu.ru

Аннотация: В статье приводится обзор исследований, посвященных явлению коммуникативной адаптации — «подстройки» собеседников друг под друга в процессе общения. В речи такая подстройка наблюдается на всех уровнях, в том числе и на фонетическом; последний и находится в центре внимания данной работы. В статье обсуждаются два основных теоретических подхода к коммуникативной адаптации: этот процесс рассматривается как автоматический или как социально обусловленный. Далее описываются стратегии адаптации: конвергенция, дивергенция, сохранение индивидуальных характеристик. Отдельный раздел посвящен методам изучения фонетических проявлений коммуникативной адаптации: способам измерения и речевому материалу, используемому в экспериментальных работах. Основная часть статьи содержит результаты исследований, полученные ведущими научными коллективами мира. Анализ данных о том, как адаптация проявляется в основных фонетических характеристиках — частоте основного тона, темпе и интенсивности речи — показывает, что из этих характеристик наиболее последовательно адаптация проявляется в интенсивности речи. Относительно других фонетических характеристик сегодня сложно составить единую картину, так как исследования разнородны с точки зрения выбора методики и материала, а полученные данные зачастую противоречивы. Все это говорит о необходимости дальнейшего изучения этого явления. В статье также обсуждается значение коммуникативной адаптации в межличностном общении, ее зависимость от различных социальных факторов (пол собеседников, роль в диалоге, социальная иерархия и др.) и связь с социальными оценками. Кроме того, демонстрируется роль адаптации в общении человека с компьютером.

Ключевые слова: адаптация, акустическая фонетика, диалог, компьютерная лингвистика, социальные факторы

Благодарности: Статья подготовлена при поддержке Российского научного фонда (грант № 19-78-10046).

Для цитирования: Качковская Т. В., Мамушина А. Д. Фонетические проявления коммуникативной адаптации в диалоге. *Вопросы языкознания*, 2021, 2: 123–141.

DOI: 10.31857/0373-658X.2021.2.123-141

Phonetic manifestations of communication accommodation in dialogue

Tatiana V. Kachkovskaia[@]

Anna D. Mamushina

Saint Petersburg State University, St. Petersburg, Russia; kachkovskaia@phonetics.pu.ru

Abstract: The paper presents a review of research on communication accommodation (also termed entrainment) — “attuning” between interlocutors over the course of interaction. In speech such “attuning” is found at all levels, including the phonetic level, which is the focus of this survey. The paper discusses

two main theoretical approaches to communication accommodation: the phenomenon is considered either an automatic or a socially mediated process. Three accommodation strategies — convergence, divergence, and maintenance — are then described. The next section considers methods employed in research on speech accommodation: metrics of accommodation and speech data used in experiments. The body of the paper provides the experimental results reported by the leading research groups. Among the phonetic parameters frequently analysed for accommodation — fundamental frequency, speaking rate, intensity — intensity shows the highest consistency in exhibiting accommodation. As for other features, it is still difficult to obtain a clear picture as the studies vary in their choice of methods and material, while the results are often controversial. Thus, the need for further research on the issue is apparent. The paper also discusses the importance of accommodation in interpersonal communication, its interplay with various social factors (gender, role in dialogue, hierarchy, etc.) and social assessments. The role of accommodation in human-computer interaction is demonstrated as well.

Keywords: acoustic phonetics, adaptation, computational linguistics, dialogue, social factors

Acknowledgements: The research is supported by Russian Science Foundation (project No. 19-78-10046).

For citation: Kachkovskaia T. V., Mamushina A. D. Phonetic manifestations of communication accommodation in dialogue. *Voprosy Jazykoznanija*, 2021, 2: 123–141.

DOI: 10.31857/0373-658X.2021.2.123-141

Введение

Речь каждого говорящего уникальна по своим акустическим характеристикам. Одна и та же фраза, произнесенная независимо двумя разными людьми, будет звучать совершенно по-разному. Однако если эта фраза прозвучит в одном диалоге с разницей в несколько реплик, то говорящий, произнося ее вслед за своим собеседником, может произнести ее очень похожим образом. Исследования на материале многих языков показывают, что своего рода «имитация» речи собеседника может наблюдаться на более глобальном уровне: говорящий часто «подстраивается» под собеседника по целому ряду акустических признаков: громкости, темпу, диапазону частоты основного тона голоса и др. Такая «подстройка» собеседников друг под друга — наряду с жестами, мимикой и пр. — получила название «коммуникативная адаптация».

Научное лингвистическое сообщество проявляет большой интерес к теме взаимной коммуникативной адаптации собеседников. В последние 10 лет эта тема становится все более востребованной в связи с развитием речевых технологий и интерфейсов общения человека с компьютерными системами. На сегодняшний день системы синтеза речи и автоматические системы ведения диалога достигли высокого уровня. В связи с этим разработчики этих технологий обратились к фонетистам с новыми исследовательскими вопросами относительно того, как сделать эти системы более похожими на живого человека. В частности, одной из актуальных задач становится подстройка голоса компьютера под собеседника-человека, позволяющая сделать диалог с компьютером более комфортным и приятным для пользователя.

Основным объектом исследований является то, как меняются акустические, сегментные и просодические характеристики речи собеседников в процессе диалога. Сознательная адаптация под диктора обычно подразумевает компенсацию социального статуса одного из собеседников или противопоставление одного собеседника другому. Бессознательная адаптация скорее оказывается стремлением к комфортному диалогу, где оба собеседника не только приходят в итоге к некому общему подязыку, но и к «единой» речевой форме. Современные фонетические исследования показывают универсальность явления фонетической адаптации, однако то, какие именно и как речевые признаки изменяются, зависит в том числе и от специфики каждого конкретного языка. В большинстве современных научных работ анализируются интенсивность речи, мелодические и темпоральные характеристики, качество голоса и др. Одновременно с этим исследуется проявление

адаптации в зависимости от пола собеседников, степени их знакомства, положения в социальной иерархии и пр.

Фонетические проявления коммуникативной адаптации изучаются на материале английского [Natale 1975; Pardo 2006; Lee et al. 2010; Levitan, Hirschberg 2011; Babel, Bulatov 2012], итальянского [Savino et al. 2016], испанского [Levitan et al. 2015b], китайского [Xia et al. 2014], немецкого [Schweitzer, Lewandowski 2013; Zellers 2016], польского [Karpiński et al. 2014], португальского [Cabarrão et al. 2016], словацкого [Beňuš et al. 2014], французского [Bailly, Lelong 2010], шведского [Edlund et al. 2009], японского [Suzuki, Katagiri 2007; Kawahara et al. 2015] языков.

Обращаясь к теме коммуникативной адаптации, отечественные лингвисты уделяли основное внимание прагматике диалогической речи. Н. Д. Арутюнова [1981] и Е. В. Падучева [1982] занимались прагматическими аспектами синтаксиса и семантики диалогической речи. Одновременно с этим Л. Р. Зиндер, Н. Д. Светозарова и их ученики занимались прагматическими аспектами интонации диалогической речи [Филиппов 1982; Соболева 1994]. С. М. Соболева [1994] на небольшом материале микродиалогов показала, что в некоторых случаях возникает «определенная интонационная зависимость прямого ответа от вопроса, проявляющаяся в тенденции сближения частоты основного тона в зоне интонационного центра ответа с частотой основного тона вопроса в аналогичном участке интонационного контура». Среди недавних работ отечественных лингвистов следует отметить краткую обзорную работу Т. И. Шевченко и И. Е. Галочкиной [2016], а также исследование А. В. Горбылевой [2019] на англоязычном материале.

В других странах эта тема сейчас очень активно развивается: разрабатываются новые методы оценки коммуникативной адаптации, сравниваются данные по разным языкам, результаты исследований внедряются в диалоговые системы. Целью данной статьи является обзор зарубежных работ, посвященных этому явлению. Здесь будут рассмотрены теоретические аспекты и методы изучения адаптации, а также обобщены результаты исследований.

1. Коммуникативная адаптация и основные подходы к ее изучению

Коммуникативная адаптация — это приспособление участников коммуникации друг к другу. Оно может проявляться в жестах, позах, выражениях лица [Dabbs 1969; Maurer, Tindall 1983; Chartrand, Bargh 1999; Hess, Blairy 2001; Shockley et al. 2004] и, конечно же, в речи.

В англоязычной литературе для обозначения данного явления используются различные термины¹: accommodation [Namy et al. 2002; De Looze et al. 2014], alignment [Pickering, Garrod 2004; Suzuki, Katagiri 2007], entrainment [Brennan 1996], convergence [Pardo 2006; Bailly, Lelong 2010], imitation [Goldinger 1998; Fowler et al. 2003], coordination [Branigan et al. 2000; Niederhoffer, Pennebaker 2002], congruence [Welkowitz et al. 1972]. В этой работе мы будем применять термин «адаптация». В русскоязычных работах также используются термины «аккомодация» и «согласованность» [Шевченко, Галочкина 2016; Горбылева 2019].

Исследования, посвященные коммуникативной адаптации, можно поделить на две большие группы. Они различаются взглядами авторов на причины коммуникативной адаптации и, как следствие, методами, используемыми в экспериментальных исследованиях.

В первом случае коммуникативная адаптация рассматривается как автоматический процесс, обусловленный связью между восприятием и производством речи [Fowler 1986; Pickering, Garrod 2004]. Предполагается, что в основе этого процесса лежит прайминг (англ. priming): так, из синонимического ряда говорящий скорее выберет ту речевую единицу,

¹ Более подробный список терминов можно найти в диссертации Й. Эдлунда [Edlund 2011: xxxiii].

которая ранее уже звучала в речи его собеседника; подобным образом говорящие могут «копировать» мелодические контуры, отдельные речевые характеристики. Таким образом, сходства в речи собеседников возникают за счет взаимной активации одних и тех же единиц, которые затем и используются далее в диалоге; адаптация происходит автоматически.

Эксперименты в рамках этого подхода используют, как правило, задания на повторение слов или фраз за диктором (англ. *shadowing task*; достаточно широкий список таких работ приведен в исследовании Дж. Пардо [Pardo et al. 2017]). Подавляющее большинство подобных работ выполнено на англоязычном материале. Эксперименты строятся следующим образом: сначала участник читает список слов с листа, а затем повторяет те же слова за диктором. После этого произнесение каждого слова из первой части (чтение) сравнивается с произнесением из второй части (повторение) посредством перцептивного эксперимента либо инструментального анализа (измеряются форманты гласных, время начала озвончения смычных согласных (англ. *voice onset time*, VOT), частота основного тона, аллофонное варьирование и т. п.). Кроме того, часто отслеживается и продолжительность эффекта адаптации: в эксперимент добавляется третье произнесение — чтение слов с листа уже после повторения за диктором.

Большой вклад в развитие этого направления внес психолог С. Голдингер [Goldinger 1998]. Для выявления подстройки он использовал АВХ-сравнение чтения и повторения (А/В) в речи участника с произнесением диктора (Х). Адаптацией считался случай, когда аудиторы выбирали повторение за диктором как более близкое к дикторскому произнесению, нежели чтение. По результатам перцептивного эксперимента более половины слушателей показали именно такую взаимосвязь, что говорит о наличии уподобления.

Вслед за С. Голдингером [Ibid.] задания на повторение применяли и другие исследователи [Fowler et al. 2003; Namy et al. 2002; Shockley et al. 2004; Pardo 2006; Pardo et al. 2013b; Babel 2012]. К. Шокли и его коллеги [Shockley et al. 2004], к примеру, в своем эксперименте использовали дикторские произнесения с искусственно увеличенным временем начала озвончения смычных; таким образом исследователи проверяли, насколько эта характеристика участвует в адаптации. Было выяснено, что при повторении за диктором данный параметр действительно оказался больше, чем в чтении; перцептивный эксперимент также подтвердил, что повторения чаще опознавались аудиторами как более близкие к образцу. Адаптация по времени начала озвончения смычных была также обнаружена в работе К. Фаулер и ее коллег [Fowler et al. 2003]. Дж. Пардо и ее коллеги [Pardo et al. 2013b] исследовали аналогичным методом частоту основного тона, спектр и длительность гласных и выяснили, что все три характеристики в совокупности связаны с воспринимаемой адаптацией, тогда как связь каждой отдельной характеристики с адаптацией статистически незначима. В другой работе [Babel 2012] также изучался спектр гласных, и наибольшая адаптация была обнаружена для гласных нижнего подъема; автор связывает это с большей вариативностью этих гласных в диалектах американского английского.

Заметным ограничением экспериментов с повторениями является их оторванность от ситуации реального общения. Поэтому некоторые работы используют записи спонтанной речи, выбирая из них одни и те же слова, произнесенные обоими участниками, и сравнивают их с помощью перцептивных экспериментов [Pardo 2006; Pardo et al. 2013a].

Тем не менее даже в сугубо лабораторных условиях заданий на повторение было обнаружено влияние социальных факторов на адаптацию. Об этом свидетельствуют, например, работы М. Бейбл и Д. Булатов [Babel, Bulatov 2012] и Л. Нами и ее коллег [Namy et al. 2002], где в экспериментах с повторениями женщины проявляли большую адаптацию, чем мужчины.

Важность социального аспекта адаптации лежит в основе второго подхода к адаптации: предполагается, что люди используют различные стратегии адаптации для достижения социальных целей. Здесь основное внимание исследователей сосредоточено на социальных факторах, влияющих на адаптацию. Объектом изучения обычно служит спонтанная речь. Центральной теорией для этого подхода является теория коммуникативной

адаптации (англ. Communication accommodation theory, сокращенно CAT), разработанная Говардом Джайлзом.

Теория коммуникативной адаптации выросла из более узкой теории — теории адаптации речи (англ. Speech accommodation theory, сокращенно SAT). Годом ее рождения можно назвать 1973 — тогда были опубликованы работы Г. Джайлза [Giles 1973; Giles et al. 1973]. В рамках этой теории рассматриваются и фонетические аспекты; при разработке этой теории Джайлз и его коллеги опирались на уже имеющиеся на тот момент данные о фонетических проявлениях адаптации. Затем стали появляться исследования адаптации в других аспектах коммуникации, например, адаптация в жестах, выражении лица и т. д., таким образом, сформировалась более широкая теория коммуникативной адаптации. Согласно теории Джайлза, основная цель адаптации — установление желаемой социальной дистанции между собеседниками. Для достижения этой цели могут использоваться три стратегии: конвергенция, дивергенция и сохранение индивидуальных характеристик (англ. convergence, divergence, maintenance).

Конвергенция (уподобление) описывает процесс, при котором речь собеседников сближается по тем или иным параметрам. К примеру, в ходе диалога оба собеседника могут со временем достигнуть близких значений по темпу речи или громкости, может уменьшаться различие дикторов по частоте основного тона. При этом характеристики голоса могут меняться как у обоих дикторов, так и только у одного из них.

При дивергенции (расподоблении) различия в речи собеседников увеличиваются. Сохранение индивидуальных характеристик подразумевает, что речь говорящего не зависит от речи его собеседников, т. е. взаимовлияния не происходит.

Конвергенция, как правило, связывается с большей вовлеченностью в беседу, с более высоким качеством общения, взаимной приязнью участников, желанием одобрения, успешностью коммуникации. Так, например, в одном из исследований Джайлза канадцев, для которых английский был основным языком общения, просили оценить собеседника, обращавшегося к ним на французском или английском языках [Giles et al. 1973]. При этом участники знали, что основным языком общения этого собеседника был французский. Результаты опроса показали, что собеседник получал более высокую оценку (воспринимался как более тактичный и стремящийся преодолеть культурный разрыв), если, по мнению участников, он прикладывал больше усилий для адаптации (например, обращаясь к ним на английский, даже если для него это представляло очевидную трудность).

Несимметричная конвергенция (уподобление только одного из говорящих своему собеседнику) может свидетельствовать о социальной иерархии. Такую закономерность продемонстрировали, к примеру, С. Грегори и С. Уэбстер [Gregory, Webster 1996]. Авторы проанализировали записи ток-шоу Ларри Кинга и обнаружили, что при общении с высокостатусными гостями большую степень адаптации проявлял ведущий, тогда как гости более низкого статуса, напротив, адаптировались к ведущему больше, чем он к ним.

Дивергенция — подчеркивание различий — ассоциируется с увеличением дистанции между собеседниками, что может быть вызвано желанием подчеркнуть свою принадлежность к другой социальной группе, другим ценностным ориентирам и т. п. Ярким и часто цитируемым примером является эксперимент, описанный в работе Джайлза, Бурхиса и Тэйлора [Giles et al. 1977]. Для исследования была набрана группа участников, изучающих валлийский язык; при этом часть из них изучала его из интереса к культуре, желания сохранить культурное наследие, тогда как другим валлийский был необходим по работе. С участниками беседовал носитель нормативного английского. Его задачей было задавать вопросы о причинах, побудивших участников учить валлийский, выражая при этом сомнение в целесообразности изучения «умирающего языка с печальным будущим». В результате было обнаружено, что люди, изучающие валлийский из реального интереса к культуре, в ответ усиливали свой валлийский акцент в английской речи. Они стремились, как предполагают авторы, подчеркнуть свою связь с ценностями, которые ставил под сомнение собеседник, и показать, что его позиция им не близка. И конвергенция, и дивергенция

могут быть охарактеризованы по нескольким параметрам, например, восходящая / нисходящая (англ. upward / downward), симметричная / асимметричная (англ. symmetrical / asymmetrical); подробнее об этих характеристиках см. работу Джайлза [Giles et al. 1991].

Поскольку научный интерес к коммуникативной адаптации изначально возник в рамках психологических и социологических исследований, измерения ее проявлений на фонетическом уровне (как и на других уровнях лингвистического анализа) выступали в первую очередь средством, необходимым для изучения механизмов порождения и восприятия речи, взаимодействия людей в обществе и др. Для фонетиста, наоборот, первостепенными оказываются именно эти частные результаты, поэтому две описанные группы исследований представляют для нас интерес в равной степени. Ниже мы будем рассматривать результаты научных работ, выполненных в рамках обоих направлений.

2. Методы изучения коммуникативной адаптации в речи

С методологической точки зрения коммуникативная адаптация измеряется разными способами². Наиболее распространенными методами являются измерение степени близости (или «сближения»; англ. proximity), синхронности (или «синхронизации»; англ. synchrony) и конвергенции (англ. convergence) [Levitan, Hirschberg 2011; Pérez et al. 2016]. Эти меры проиллюстрированы на рис. 1.

Степень близости — это статичная мера, характеризующая разницу значений того или иного параметра в речи собеседников во всем диалоге. Эта мера часто используется, если один и тот же диктор участвовал в диалогах с несколькими разными собеседниками: в таком случае вычисляют среднее значение какого-либо параметра по каждому диалогу, а затем сравнивают, как меняется это значение — как правило, с учетом социальных различий между собеседниками.

Синхронность и конвергенция измеряются в динамике. Под конвергенцией понимается постепенное уменьшение разницы тех или иных показателей в речи собеседников. Например, если человек стал говорить медленнее, приспосабливаясь к более медленному темпу речи своего собеседника, это уменьшение разницы считается конвергенцией. Синхронность — это более локальная мера, характеризующая синхронное изменение показателей в одном направлении. То есть если человек начинает говорить громче в ответ на возрастающую громкость речи собеседника и тише в ответ на снижение громкости — это описывается как синхронность.

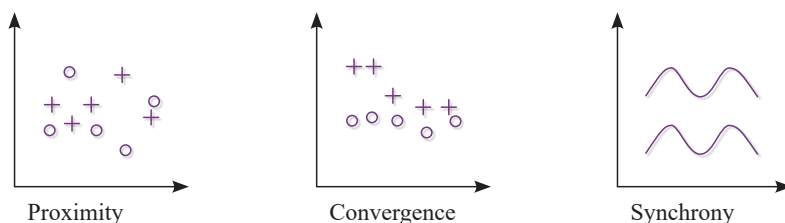


Рис. 1. Меры адаптации: степень близости, конвергенция, синхронность

По оси X отобразено время, по оси Y — значение параметра. Круги и знаки плюс на первых двух изображениях показывают значения в речи двух собеседников. Источник: [Levitan, Hirschberg 2011: 3081].

² Не стоит путать способы измерения со стратегиями адаптации, о которых говорилось в конце предыдущего раздела.

Описанные меры могут быть и отрицательными, то есть отражать изменения параметра в противоположном от собеседника направлении. Противоположностью конвергенции является дивергенция (англ. *divergence*), противоположностью синхронности — асинхронность (англ. *asynchrony*) или отрицательная синхронность (*negative synchrony*) (см. рис. 2).

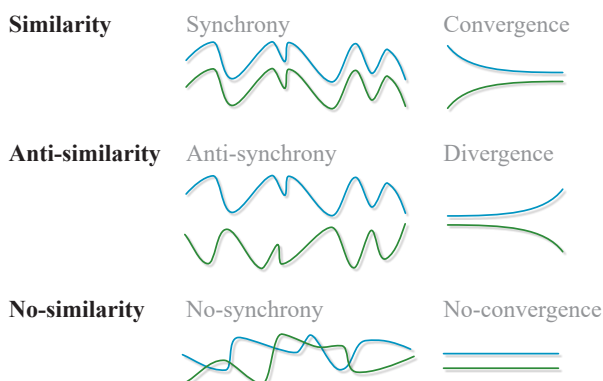


Рис. 2. Меры адаптации: положительные и отрицательные

По рядам слева направо: 1) уподобление: синхронность, конвергенция; 2) расподобление: асинхронность, дивергенция; 3) взаимонезависимость: отсутствие синхронности, отсутствие конвергенции. Источник: [De Looze, Rauzy 2011: 1393].

Выбор материала исследования зависит от того, в рамках какого подхода исследуется коммуникативная адаптация. При трактовке адаптации как автоматического процесса используется лабораторный материал — задания на повторение слов/фраз за диктором. При изучении адаптации как социально обусловленного процесса материалом обычно служит спонтанная речь. Это может быть общение на свободные темы [Natale 1975; Ireland et al. 2011; Schweitzer, Lewandowski 2013; De Looze et al. 2014], на заданную тему [Ward, Litman 2007; Lee et al. 2010; Kawahara et al. 2015; Weidman et al. 2016] или выполнение каких-либо заданий. Как правило, это задания, в которых у собеседников есть одна общая цель; предполагается, что решение общей задачи способствует подстройке собеседников друг под друга. Очень распространены такие задания, как, поиск одинаковых картинок/элементов на картинках и расположение картинок в нужном порядке [Nenkova et al. 2008; Xia et al. 2014; Beňuš et al. 2014; Levitan et al. 2015b; Pérez et al. 2016], ранжирование предметов для выживания в опасной ситуации [Kousidis et al. 2009] и т. д. Наиболее популярны задания с картами, где одному из собеседников необходимо начертить на карте маршрут по инструкции другого [Pardo 2006; Reitter et al. 2006; Levitan, Hirschberg 2011; Pardo et al. 2013a; Cabarrão et al. 2016; 2018]; пример задания с маршрутом приведен на рис. 3 (с. 130).

Среди известных корпусов, используемых для изучения адаптации, можно назвать Columbia Games Corpus для американского английского [Gravano 2009], записанные по его образцу Tongji Games Corpus для китайского [Xia et al. 2014], Porteno Spanish Games Corpus для испанского (Аргентина) [Levitan et al. 2015a] и Sk-Games Corpus [Beňuš 2012] для словацкого, кроме того, корпус CORAL для европейского португальского [Trancoso et al. 1998], а также HCRC Map Task Corpus [Anderson et al. 1991] и созданный по его образцу новый Montclair Map Task Corpus [Pardo et al. 2019] (американский английский).

Коммуникативная адаптация в речи изучается на различных уровнях. Результаты исследований показывают, что адаптация проявляется на лексическом [Brennan 1996; Pickering, Garrod 2004; Ward, Litman 2007; Nenkova et al. 2008], грамматическом (в том числе в выборе синтаксических конструкций) [Branigan et al. 2000; Niederhoffer, Pennebaker 2002; Cleland, Pickering 2003; Haywood et al. 2005; Reitter et al. 2006; Ireland et al. 2011] и фонетическом уровне.

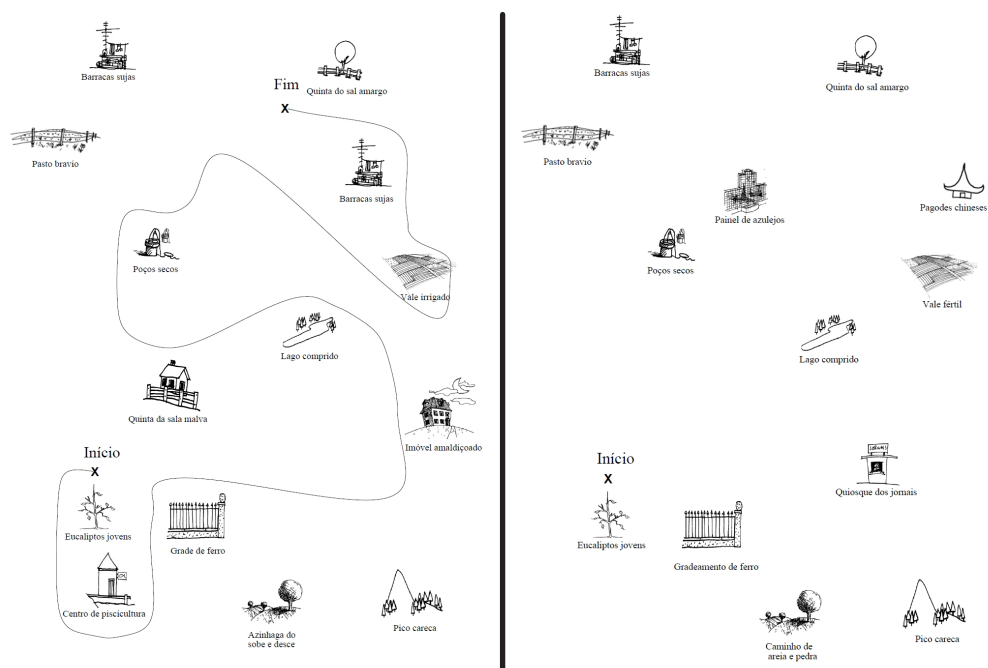


Рис. 3. Пример задания с маршрутом

Слева: карта инструктирующего (его задача — описать собеседнику нарисованный маршрут), справа: карта инструктируемого (его задача — начертить маршрут, следуя указаниям собеседника). Источник: [Trancoso et al. 1998: 103–104].

На фонетическом уровне изучаются такие параметры, как:

- интенсивность речи [Natale 1975; Ward, Litman 2007; Kousidis et al. 2009; Levitan, Hirschberg 2011; Vaughan 2011; De Looze et al. 2014; Xia et al. 2014; Beňuš et al. 2014; Levitan et al. 2015a; 2015b; Székely et al. 2015; Pérez et al. 2016; Savino et al. 2016],
- частота основного тона [Ward, Litman 2007; Kousidis et al. 2009; De Looze, Rauzy 2011; Levitan, Hirschberg 2011; Vaughan 2011; Babel, Bulatov 2012; Pardo et al. 2013b; De Looze et al. 2014; Xia et al. 2014; Beňuš et al. 2014; Levitan et al. 2015a; 2015b; Cabarrão et al. 2016; Pérez et al. 2016; Savino et al. 2016; Zellers 2016; Pardo et al. 2017],
- в том числе интонационные конструкции [Reichel, Cole 2016],
- темп [Kousidis et al. 2009; Pardo et al. 2013a; Schweitzer, Lewandowski 2013; De Looze et al. 2014; Karpinski et al. 2014; Xia et al. 2014; Beňuš et al. 2014; Cabarrão et al. 2016; Pérez et al. 2016; Savino et al. 2016; Zellers 2016],
- длительность высказываний [Ray, Webb 1966; Matarazzo, Wiens 1967; Pardo et al. 2013a; Cabarrão et al. 2016],
- паузы между высказываниями и паузы хезитации [Beňuš et al. 2012; Pardo et al. 2013a],
- характеристики качества голоса [Levitan, Hirschberg 2011; Vaughan 2011; Levitan et al. 2015a; 2015b; Cabarrão et al. 2016; Pérez et al. 2016],
- форманты гласных [Delvaux, Soquet 2007; Babel, Bulatov 2012; Pardo et al. 2013b; 2017],
- время начала озвончения смычных согласных (англ. voice onset time, VOT) и т. д.

Иногда исследуется один показатель, но нередко рассматривается сразу несколько. Используются и наборы акустических признаков, например GeMaps [Eyben et al. 2016].

3. Проявления коммуникативной адаптации в основных акустических характеристиках речи

Несмотря на немалое количество исследований, посвященных адаптации, составить некую единую картину по их результатам довольно сложно. В разных работах используются разные методы, изучаются разные параметры и в разных условиях, поэтому нередко полученные результаты оказываются несравнимы. Сложность заключается и в самом явлении адаптации: на нее может влиять огромное количество факторов, учесть влияние всех практически невозможно. Даже в рамках одного эксперимента один и тот же параметр может показать в одном случае конвергенцию, в другом — дивергенцию, в третьем — ни то ни другое. В одном диалоге может наблюдаться конвергенция по одним параметрам и дивергенция по другим [Pardo et al. 2013a; Levitan et al. 2015a; Savino et al. 2016].

Далее будут рассмотрены часто изучаемые на предмет адаптации характеристики и данные, полученные для этих параметров в ходе исследований.

Интенсивность речи является характеристикой, которая чаще прочих показывает положительные результаты в исследованиях адаптации. Среди ярких примеров можно назвать широко цитируемую работу М. Натале [Natale 1975], где громкость голоса одного из собеседников в диалоге искусственно повышалась и понижалась, а второй собеседник в ответ менял громкость своего голоса в том же направлении.

Интересное свойство адаптации по интенсивности обнаружилось в исследовании Ш. Бенюша и его коллег [Beňuš et al. 2014]. В данном эксперименте сравнивалась речь двух собеседников, а также речь участника диалога с его собственной речью, но из другого диалога. Интенсивность оказалась единственной характеристикой из всех рассмотренных, которая показала большее сходство при сравнении двух собеседников, чем при сравнении диктора с самим собой. Кроме этого, интенсивность в целом была показателем, который чаще всего давал положительный результат при применении разных мер адаптации. Аналогичный результат был получен и в ряде других работ [Levitan, Hirschberg 2011; Xia et al. 2014; Levitan et al. 2015a; Weidman et al. 2016]. В работе С. Коусидиса и его коллег [Kousidis et al. 2009] интенсивность речи (а также частота основного тона) показали наиболее высокие коэффициенты корреляции в речи собеседников. Тем не менее распознавание по данному параметру также обнаруживается [Levitan et al. 2015a; Pérez et al. 2016]. У. Райхель и его коллеги [Reichel et al. 2018], в работе которых наблюдался такой результат, предполагают, что взаимное повышение громкости голоса может быть связано с борьбой за право говорить. Кроме того, как показывает эксперимент Е. Секей [Székely et al. 2015], адаптация по интенсивности зависит от условий коммуникации. В тишине она проявляется сильнее, чем в шуме; более того, если человек при шуме говорит тихим голосом, то его собеседник повышает громкость своего голоса (возникает отрицательная синхронность). По мнению авторов исследования, это один из способов побудить собеседника говорить громче, не прося его об этом явно.

Частота основного тона (ЧОТ) — акустическая характеристика, воспринимаемая как высота голоса — также часто измеряется в исследованиях адаптации, однако касающиеся ее результаты несколько менее однозначны. Так, например, в упомянутой работе Коусидиса и его коллег [Kousidis et al. 2009] для ЧОТ наблюдались высокие коэффициенты корреляции в речи собеседников. В двух других исследованиях [Levitan, Hirschberg 2011; Xia et al. 2014] также была выявлена адаптация по ЧОТ при применении нескольких разных мер адаптации. Интересно, что в последней из упомянутых работ, посвященной китайскому языку, отмечались достаточно высокие коэффициенты синхронности, тогда как в работе Р. Левитан и ее коллег [Levitan et al. 2015a] в китайском языке была обнаружена отрицательная синхронность (авторы связывают это с тем, что китайский является тоновым языком, и ЧОТ в нем играет более сложную роль). Некоторые исследования свидетельствуют о дивергенции по ЧОТ [Pérez et al. 2016; Zellers 2016]. В эксперименте М. Савино

и ее коллег [Savino et al. 2016] в одном из диалогов наблюдалась конвергенция по диапазону ЧОТ и дивергенция по средней ЧОТ. Другое исследование [Ward, Litman 2007] выявило дивергенцию по минимальному значению ЧОТ и отсутствие адаптации по максимальному и среднему значениям.

Еще более противоречивые данные касаются **темпа**: этот параметр в одних случаях показывает отрицательный результат, в других — положительный [Xia et al. 2014; Savino et al. 2016]. В ряде исследований [Kousidis et al. 2009; Levitan, Hirschberg 2011] обнаруживается адаптация по темпу, но с меньшим эффектом, нежели по ЧОТ и интенсивности. Однако некоторые исследования говорят об отсутствии конвергенции по темпу [Zellers 2016], другие — о наличии дивергенции [Schweitzer, Lewandowski 2013]. В эксперименте М. Карпинского и его коллег [Karpiński et al. 2014] корреляция по темпу, хотя и наблюдалась нечасто, была достаточно высокой. Авторы предполагают, что адаптация по темпу может быть локальным явлением, возникающим на тех участках диалога, где общение происходит наиболее свободно и непринужденно. Кроме того, есть данные о том, что адаптация по темпу зависит от роли собеседника в диалоге [Pardo et al. 2013a] (ведущий-ведомый). Сравнение результатов осложняется также тем, что в разных работах используются разные методы измерения темпа: количество гласных в минуту [Kousidis et al. 2009], слогов в секунду [Levitan, Hirschberg 2011; Xia et al. 2014; Savino et al. 2016; Zellers 2016], слогов в высказывании [Schweitzer, Lewandowski 2013], слов в секунду [Pardo et al. 2013a].

Помимо акустических характеристик, при исследовании коммуникативной адаптации анализируются также различные элементы организации диалога, такие как паузы (в том числе паузы hesitation), контактоподдерживающие реплики³ (например, «угу»), смех, собственно реплики диалога (обычно анализируется их длина).

Одной из ранних и часто цитируемых работ является статья Дж. Матараццо и А. Винза [Matarazzo, Wiens 1967]. Ее авторы продемонстрировали, что увеличение / уменьшение длительности высказывания в речи говорящего, выступающего в роли ведущего (берущего интервью у своего собеседника), вызывает соответствующее изменение длительности высказывания в речи собеседника (интервьюируемого).

Интерес исследователей также привлекают контактоподдерживающие реплики (например, «угу» и «ага») и связанные с ними явления. Так, в работе Р. Левитан [Levitan 2014] был проведен достаточно подробный анализ просодических маркеров, с помощью которых говорящий «призывает» собеседника отреагировать на сообщение (англ. backchannel-inviting cues) — то есть произнести контактоподдерживающую реплику. (Среди таких маркеров отмечают восходящую интонацию в конце высказывания, более высокий уровень ЧОТ, повышение громкости, изменение качества голоса, большую продолжительность речи от паузы до паузы и др.) По результатам эксперимента было выяснено, что собеседники действительно используют схожий набор таких сигналов.

Исследование М. Хельднера и его коллег [Heldner et al. 2010] показало, что в контактоподдерживающих репликах ярко проявляется адаптация по ЧОТ. Если сравнивать разницу в ЧОТ между предыдущей репликой одного собеседника и следующей за ней репликой второго, то контактоподдерживающие реплики обычно более близки к предыдущему высказыванию, чем прочие реплики к предшествующим им высказываниям. Авторы предполагают, что это связано с «фоновым» характером таких реплик: чем меньше отличается контактоподдерживающая реплика от предыдущего высказывания собеседника, тем меньше она выделяется. Кроме того, есть сведения об адаптации по типу заполнителей пауз: так, в работе У. Коэн Прива и Ч. Сэнкер [Cohen Priva, Sanker 2018] была обнаружена

³ Здесь и далее термином «контактоподдерживающая реплика» обозначается реплика, не приводящая к передаче слова в диалоге и показывающая, что собеседник слушает и понимает, что ему говорят (англ. backchannel). В русскоязычной литературе используются и другие термины, в частности — «реплика-реакция», «реактивная реплика», «реплика обратной связи».

адаптация по соотношению количества «um» и «uh» (двух самых частотных видов заполнителей пауз в англоязычной речи). Иными словами, в ходе диалога собеседники часто неосознанно «заимствуют» заполнители пауз друг у друга.

В эксперименте Р. Левитан и ее коллег [Levitan et al. 2015b] исследовалась разница в распределении способов передачи реплики (в зависимости от того, прервал ли говорящий своего собеседника, начал ли говорить после окончания его высказывания, произошло ли наложение речи собеседников или была произнесена контактоподдерживающая реплика). С помощью сравнения реальных и искусственно созданных случайных диалогов было обнаружено, что соотношение разных типов реплик в речи реальных собеседников сходно.

В другом исследовании [Bilous, Krauss 1988] изучалась адаптация в речи мужчин и женщин по нескольким параметрам организации диалога. В речи женщин была обнаружена конвергенция по таким характеристикам, как общее количество слов, длительность высказываний, количество прерываний одного собеседника другим, частота употребления пауз, и дивергенция по частоте употребления контактоподдерживающих реплик и частоте появления смеха. В речи мужчин конвергенцию показали такие параметры, как длительность высказываний, частота употребления пауз и контактоподдерживающих реплик, смех. Эти результаты позволили авторам сделать ряд выводов о ведении диалога мужчинами и женщинами. В частности, на основании данных о смене реплик было показано, что гипотезу о доминировании мужчин в диалоге, широко обсуждавшуюся в то время, необходимо переформулировать: доминирование состоит не в изменении речевых характеристик, а наоборот, в сохранении их неизменными.

Тем не менее, как показывает работа Дж. Пардо [Pardo et al. 2013a], такие характеристики, как объем речи и частота употребления заполненных пауз, зависят от роли говорящего в диалоге (ведущий-ведомый).

4. Факторы, влияющие на коммуникативную адаптацию

Среди факторов, часто рассматриваемых с точки зрения их влияния на адаптацию, можно назвать пол собеседников, их роли в диалоге, а также такие социальные характеристики, как, например, социальная желательность (social desirability), проявление эмпатии, социальная иерархия.

Что касается **пола** собеседников, результаты исследований здесь довольно противоречивы. Например, в эксперименте М. Карпинского и его коллег [Karpinski et al. 2014] адаптация чаще наблюдалась в диалогах между собеседниками разного пола, а в работе В. Кабарран [Cabarrão et al. 2016] — в диалогах между собеседниками одного пола, особенно женского. В другой работе [Xia et al. 2014] адаптация по наибольшему количеству параметров была характерна для собеседников разного пола, по наименьшему — для собеседников-мужчин. В уже упомянутой работе Ф. Билэс и Р. Краусса [Bilous, Krauss 1988] женщины проявляли конвергенцию и дивергенцию по нескольким характеристикам, мужчины — только конвергенцию, при этом поведение участников достаточно сильно отличалось не только в зависимости от пола говорящего, но и в зависимости от пола собеседника. Так, женщины выбирали доминирующую модель поведения только в диалогах с женщинами. В исследовании Дж. Пардо [Pardo 2006] было выявлено взаимодействие факторов пола и роли в диалоге: так, для женщин было обнаружено, что собеседники, дающие инструкции (в задании с объяснением маршрута), больше подстраивались под слушающего; у мужчин же наблюдалась обратная взаимосвязь; при этом в целом у женщин наблюдалось меньше конвергенции. Автор отмечает, что такие результаты свидетельствуют о различной интерпретации мужчинами и женщинами заданных ролей в диалоге. Такой же результат

был получен в эксперименте У. Райхеля [Reichel et al. 2018]⁴: у женщин наблюдалась большая степень адаптации, особенно в роли инструктирующего; мужчины же, наоборот, подстраивались меньше и в целом, и в роли инструктирующего особенно. Так как задание заключалось в достижении общей цели, авторы делают вывод, что мужчины и женщины используют различные стратегии для успешного выполнения поставленной задачи: мужчины склонны подчеркивать иерархию с помощью расподобления, при этом инструктируемый адаптируется больше, признавая ведущую позицию инструктирующего, тогда как женщины стараются установить взаимопонимание путем адаптации, и ответственность за это возлагается в большей степени на инструктирующего. Кроме того, авторы отмечают, что в случае ЧОТ уподобление заключается в основном в синхронности, то есть дикторы разного пола не стремятся прийти к общему среднему значению, а синхронизируют использование более высокого / низкого для своей речи диапазона. Кроме того, есть сведения о том, что ЧОТ зависит от степени привлекательности собеседника: в некоторых исследованиях и мужчины, и женщины увеличивают диапазон ЧОТ при общении с более привлекательным для них собеседником и повышают ЧОТ, по другим данным — понижают [Michalsky, Schoormann 2017]. Учитывая расхождения в результатах исследований влияния пола собеседников на степень адаптации, можно сделать следующий вывод: если пол и влияет на адаптацию, то не напрямую, а как часть комплексного взаимодействия ряда социальных и индивидуальных факторов. К аналогичным выводам пришли А. Вайзе и его коллеги в одной из недавних работ, где анализировалась речь носителей английского и китайского языков, беседующих на английском языке [Weise et al. 2019]. Их данные показали, что степень адаптации не зависит от пола, родного языка или комбинации этих факторов.

Более поздняя работа Дж. Пардо и ее коллег [Pardo et al. 2013a] была целиком посвящена влиянию фактора **роли** в диалоге на адаптацию (ведущий–ведомый, или инструктирующий–инструктируемый). Результаты их исследования показали необходимость учета изначально заданной роли при исследовании адаптации и взаимодействие этого фактора с фактором пола собеседников. Однако выводы о том, в какой роли говорящий адаптируется больше, сделать трудно, поскольку результаты этой работы противоречат полученным ранее результатам [Pardo 2006].

Адаптация зависит и от цели общения. Так, в работе У. Райхеля и Дж. Коул [Reichel, Cole 2016], где исследовались мелодические контуры⁵, было обнаружено, что адаптация проявляется сильнее при выполнении заданий, требующих сотрудничества, чем в условиях соревнования. Кроме того, как уже упоминалось, влиять на адаптацию могут степень вовлеченности, взаимопонимания, согласия, привлекательности. В работе М. Натале [Natale 1975] больше конвергенции наблюдалось в речи участников с высоким показателем социальной желательности⁶, в исследовании М. Савино и ее коллег [Savino et al. 2016] — в речи участников с большим уровнем эмпатии. В эксперименте А. Швейцер и Н. Левандовски [Schweitzer, Lewandowski 2013] было показано, что при высокой взаимной привлекательности собеседников степень дивергенции оказывается ниже. Тем не менее есть и другие выводы: например, по данным исследования Х. Переса и его коллег [Pérez et al. 2016], дивергенция наблюдается при хорошем взаимопонимании и продуктивном взаимодействии участников. Однако следует заметить, что в приведенных работах рассматривались разные акустические параметры. В другом исследовании [De Looze et al. 2014] изучалась связь адаптации по различным акустическим характеристикам с социальными оценками (такими

⁴ В данной работе использовались задания на расположение картинки в нужном месте на экране; роль инструктирующего заключалась в описании положения картинки.

⁵ Набор стилизованных контуров, полученных с помощью инструмента CoPaSul [Reichel 2014].

⁶ Измерялась с помощью шкалы социальной желательности Кроуна — Марлоу (англ. Marlowe — Crowne Social Desirability Scale) — теста, направленного на выявление потребности в одобрении [Crowne, Marlowe 1960]. Данный опросник состоит из 33 утверждений, с которыми испытуемый должен выразить согласие или несогласие.

как естественность диалога, вовлеченность, взаимопонимание, взаимная привлекательность собеседников, неуравновешенность диалога, склонность перебивать друг друга), выявленными через перцептивный эксперимент. Некоторые параметры действительно обнаружили такую связь (например, синхронность по ЧОТ и такие оценки, как естественность беседы, вовлеченность и симпатия собеседников, асинхронность по ЧОТ и интенсивности и неуравновешенность диалога), другие — нет (например, темп). Таким образом, обнаруженные взаимосвязи в целом соответствуют предположениям теории Джайлза, однако взаимодействие конкретных характеристик с социальными факторами требует уточнения.

Влияние социальных и индивидуальных факторов на коммуникативную адаптацию продолжает активно исследоваться. В одной из последних работ Н. Левандовски и М. Йилка [Lewandowski, Jilka 2019] показано, что на степень адаптации влияют способности говорящего к фонетике (англ. phonetic talent). Кроме того, в этой публикации показано влияние ряда когнитивных и личностных особенностей говорящего на фонетическую адаптацию. Обобщая результаты этого и предыдущих исследований, авторы приводят схему факторов, влияющих на адаптацию (см. рис. 4): базовые языковые навыки, способности к фонетике, личные качества говорящего, когнитивные характеристики (внимание, память), необходимость в общественном одобрении, восприятие собеседника (привлекательность, «дружественность», симпатия).

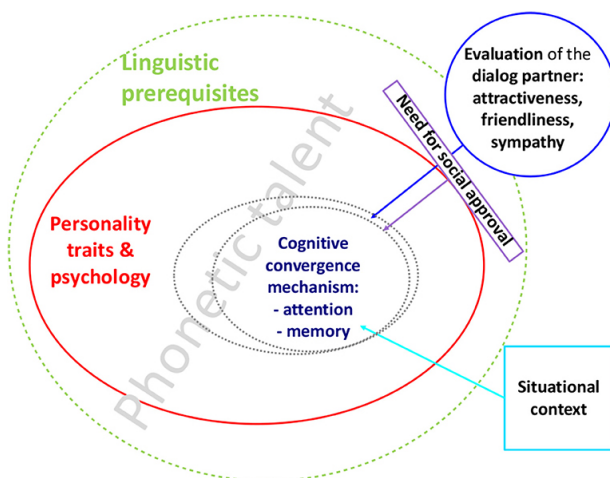


Рис. 4. Факторы, влияющие на фонетическую адаптацию
Источник: [Lewandowski, Jilka 2019: 3].

5. Коммуникативная адаптация в общении человека и компьютера

Сведения о коммуникативной адаптации в речи широко применяются в технологиях, обеспечивающих общение человека с компьютером (англ. human-computer interaction). Основной вектор развития данных технологий направлен на то, чтобы сделать поведение компьютера как можно более похожим на поведение человека. Поскольку при естественном общении люди адаптируют свою речь к речи друг друга, компьютеру также необходимо уметь это делать. Заложенная в системы синтеза речи функция подстройки (подробно о создании таких систем см. [Levitani 2014]) может сделать их более удобными

и приятными для пользователя. На сегодняшний день уже доказана высокая эффективность систем искусственного порождения речи, способных к такому уподоблению. Так, пользователи давали более высокие оценки системам, которые осуществляли подстройку по темпу [Ward, Nakagawa 2002], по частоте основного тона [Lubold et al. 2015], по темпу, громкости и диапазону частоты основного тона [Nass, Lee 2001]. Однако результат может быть и обратным [Levitani et al. 2016], что свидетельствует о необходимости дальнейшего изучения данного вопроса.

Помимо этого, исследования показывают, что человек проявляет адаптацию не только в общении с другими людьми, но и в общении с компьютером. Это можно наиболее наглядно показать на примере лексической подстройки [Brennan 1996], когда пользователь принимает и неоднократно повторяет термин, предложенный системой. На фонетическом уровне это может быть адаптация по громкости голоса [Suzuki, Katagiri 2007; Coulston et al. 2002; Oviatt et al. 2004], темпу [Bell et al. 2003], длительности пауз [Suzuki, Katagiri 2007; Oviatt et al. 2004] и их количеству [Coulston et al. 2002]. Из этих наблюдений следует интересный вывод: возможно, зная о механизмах коммуникативной адаптации, можно в некоторой степени корректировать речь пользователя [Zoltan-Ford 1991; Brennan 1996], чтобы сделать ее более удобной для компьютерной обработки, в частности, для автоматического распознавания.

Заключение

В последнее время активно исследуется влияние ситуативных факторов на речь конкретного человека: сравниваются различные стили речи, исследуются фонетические проявления эмоций и усталости, анализируется речь в условиях шума и т. д. Коммуникативная адаптация является еще одним ситуативным фактором, ведь практически любой речевой материал представляет собой диалог в той или иной мере — а значит, фонетически этот материал может быть оформлен очень по-разному в зависимости от того, с кем говорит человек. Речевой портрет говорящего, таким образом, выступает как совокупность всех вариантов речевого поведения человека во всех возможных ситуациях общения; его невозможно описать одним набором фонетических характеристик. При анализе результатов, полученных на любом речевом материале, важно знать, какие фонетические характеристики в большей степени подвержены влиянию ситуации общения, а какие — в меньшей степени. Только в таком случае исследователь может корректно обобщить полученные данные и применить выводы к речи в целом.

На сегодняшний день опубликовано огромное количество научных работ, посвященных речевым проявлениям коммуникативной адаптации. И хотя интуитивно это явление кажется достаточно понятным — собеседники либо «уподобляются» друг другу, либо наблюдается «расподобление», либо они сохраняют свою речь неизменной — совсем не просто обнаружить, в каких акустических характеристиках это проявляется. Из всех акустических характеристик только интенсивность речи последовательно отражает коммуникативную адаптацию собеседников в разных исследованиях на материале разных языков. Можно было предполагать, что адаптация проявится в частоте основного тона — характеристике, которую говорящий легко может менять; но уже здесь не удастся выстроить четкой картины: данные в разных исследованиях противоречат друг другу. Похожая ситуация наблюдается и при анализе социальных факторов, влияющих на коммуникативную адаптацию. Существуют данные о том, что в условиях социальной иерархии «нижестоящий» подстраивается под «вышестоящего»; однако в игровых заданиях с делением на роли (инструктирующий–инструктируемый) подстраиваться под собеседника может любой из говорящих. В первых исследованиях о влиянии пола собеседников на адаптацию, казалось, выстраивалась стройная картина, основанная на доминирующей роли мужчин; однако более современные данные опровергли этот вывод — возможно, из-за смены

отношения между мужчинами и женщинами в современном обществе, а возможно — благодаря привлечению большего объема материала. Статистические модели, учитывающие несколько факторов одновременно, доказывают влияние факторов пола, роли, личных характеристик говорящих и отношений между говорящими — в этом сходятся большая часть современных исследований. Однако проследить, как влияют эти факторы по отдельности, пока не удастся, поскольку трудно создать такой материал, в котором варьировался бы только один из факторов. Очевидно, мы имеем дело со сложным взаимодействием группы социальных и индивидуальных факторов, а значит, исследования в этом направлении необходимо продолжать, в том числе и на русскоязычном материале. Это позволит научному сообществу ответить на вопрос об универсальности коммуникативной адаптации и ее проявлений в речи.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Арутюнова 1981 — Арутюнова Н. Д. Фактор адресата. *Известия АН СССР. Серия литературы и языка*, 1981, 40(4): 356–367. [Arutyunova N. D. The factor of addressee. *Izvestiya AN SSSR. Seriya literatury i yazyka*, 1981, 40(4): 356–367.]
- Горбылева 2019 — Горбылева А. В. Просодическая конвергенция как результат взаимодействия собеседников в спонтанном диалоге. *Теоретическая и прикладная лингвистика*, 2019, 5(3): 41–52. [Gorbyleva A. V. Prosodic convergence as a result of speakers' interaction in spontaneous dialogue. *Teoreticheskaya i prikladnaya lingvistika*, 2019, 5(3): 41–52.]
- Падучева 1982 — Падучева Е. В. Прагматические аспекты связности диалога. *Известия АН СССР. Серия литературы и языка*, 1982, 41(4): 305–313. [Paducheva E. V. Pragmatic aspects of dialogue cohesion. *Izvestiya AN SSSR. Seriya literatury i yazyka*, 1982, 41(4): 305–313.]
- Соболева 1994 — Соболева С. М. *Прагматический аспект интонационного взаимодействия реплик диалога*. Автореф. дис. ... канд. филол. наук. СПб.: СПбГУ, 1994. [Soboleva S. M. *Pragmatischekii aspekt intonatsionnogo vzaimodeistviya replik dialoga* [Pragmatic aspect of intonational interaction in dialogue turns]. Ph.D. diss. abstract. St. Petersburg: St. Petersburg State Univ., 1994.]
- Филиппов 1982 — Филиппов К. А. *Интонационные характеристики диалогического единства (на материале немецкого языка)*. Автореф. дис. ... канд. филол. наук. Л.: ЛГУ, 1982. [Filipov K. A. *Intonatsionnye kharakteristiki dialogicheskogo edinstva (na materiale nemetskogo yazyka)* [Prosodic features of dialogue (based on German material)]. Ph.D. diss. abstract. Leningrad: Leningrad State Univ., 1982.]
- Шевченко, Галочкина 2016 — Шевченко Т. И., Галочкина И. Е. Навстречу согласию: фонетическая адаптация собеседников в диалоге. *Вестник МГЛУ. Гуманитарные науки*, 2016, 1(740): 185–203. [Shevchenko T. I., Galochkina I. E. Ensuring convergence: Phonetic adaptation of partners in dialogue. *Vestnik MGLU. Gumanitarnye nauki*, 2016, 1(740): 185–203.]
- Anderson et al. 1991 — Anderson A. H., Bader M., Bard E. G., Boyle E., Doherty G., Garrod S., Isard S., Kowtko J., McAllister J., Miller J., Sotillo C., Thompson H. S., Weinert R. The HCRC Map Task Corpus. *Language and Speech*, 1991, 34: 351–366.
- Babel 2012 — Babel M. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, 2012, 40: 177–189.
- Babel, Bulatov 2012 — Babel M., Bulatov D. The role of fundamental frequency in phonetic accommodation. *Language and Speech*, 2012, 55: 231–248.
- Bailly, Lelong 2010 — Bailly G., Lelong A. Speech dominoes and phonetic convergence. *Proc. of Interspeech 2010*. Kobayashi T., Hirose K., Nakamura S. (eds.). International Speech Communication Association, 2010, 1153–1156. https://www.isca-speech.org/archive/interspeech_2010/i10_1153.html.
- Bell et al. 2003 — Bell L., Gustafson J., Heldner M. Prosodic adaptation in human-computer interaction. *15th International Congress of Phonetic Sciences*. Solé M. J., Recasens D., Romero J. (eds.). International Phonetic Association, 2003, 2453–2456. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/p15_2453.html.
- Beňuš 2012 — Beňuš Š. Prosodic forms and pragmatic meanings: The case of the discourse marker 'no' in Slovak. *3rd IEEE International Conf. on Cognitive Infocommunications (CogInfoCom 2012)*. Institute of Electrical and Electronics Engineers, 2012, 77–81. <https://ieeexplore.ieee.org/document/6421961>.

- Beňuš et al. 2012 — Beňuš Š., Levitan R., Hirschberg J. Entrainment in spontaneous speech: The case of filled pauses in supreme court hearings. *3rd IEEE International Conf. on Cognitive Infocommunications (CogInfoCom)*. Institute of Electrical and Electronics Engineers, 2012, 793–797. <https://ieeexplore.ieee.org/document/6421959>.
- Beňuš et al. 2014 — Beňuš Š., Levitan R., Hirschberg J., Gravano A., Darjaa S. Entrainment in Slovak collaborative dialogues. *5th IEEE International Conf. on Cognitive Infocommunications (CogInfoCom)*. Institute of Electrical and Electronics Engineers, 2014, 309–313. <https://ieeexplore.ieee.org/document/7020468>.
- Bilous, Krauss 1988 — Bilous F. R., Krauss R. M. Dominance and accommodation in the conversational behaviours of same- and mixed-gender dyads. *Language & Communication*, 1988, 8(3/4): 183–194.
- Branigan et al. 2000 — Branigan H. P., Pickering M. J., Cleland A. A. Syntactic co-ordination in dialogue. *Cognition*, 2000, 75(2): B13–25.
- Brennan 1996 — Brennan S. E. Lexical entrainment in spontaneous dialog. *Proc. of International Symposium on Spoken Dialogue (ISSD 96)*. Fujisaki H. (ed.). Tokyo: Acoustical Society of Japan, 1996, 41–44.
- Cabarrão et al. 2016 — Cabarrão V., Trancoso I., Mata A., Moniz H., Batista F. Global analysis of entrainment in dialogues. *Advances in speech and language technologies for Iberian languages. IberSPEECH 2016*. Abad A. et al. (eds.). (Lecture Notes in Computer Science, 10077.) Lisbon: Springer; Cham, 2016, 215–223. https://link.springer.com/chapter/10.1007/978-3-319-49169-1_21.
- Cabarrão et al. 2018 — Cabarrão V., Batista F., Moniz H., Trancoso I., Mata A. I. Acoustic-prosodic entrainment in structural metadata events. *Proc. of Interspeech 2018*. Yegnanarayana B. (ed.). International Speech Communication Association, 2018, 2176–2180. https://www.isca-speech.org/archive/Interspeech_2018/abstracts/2366.html.
- Chartrand, Bargh 1999 — Chartrand T. L., Bargh J. A. The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 1999, 76(6): 893–910.
- Cleland, Pickering 2003 — Cleland A. A., Pickering M. J. The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, 2003, 49(2): 214–230.
- Cohen Priva, Sanker 2018 — Cohen Priva U., Sanker C. Distinct behaviors in convergence across measures. *Proc. of the 40th Annual Meeting of the Cognitive Science Society*. Kalish C., Rau M., Zhu J., Rogers T. (eds.). Madison: Cognitive Science Society, 2018, 1515–1520.
- Coulston et al. 2002 — Coulston R., Oviatt Sh. L., Darves C. Amplitude convergence in children's conversational speech with animated personas. *Proc. of International Conf. on Spoken Language Processing (ICSLP 2002)*. Hansen J., Pellom B. (eds.). Denver: Casual Prod. Ltd., 2002, 2689–2692.
- Crowne, Marlowe 1960 — Crowne D. P., Marlowe D. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 1960, 24(4): 349–354.
- Dabbs 1969 — Dabbs J. M., Jr. Similarity of gestures and interpersonal influence. *Proc. of the 77th Annual Convention of the American Psychological Association*. Washington: American Psychological Association, 1969, 4(1): 337–338.
- De Looze, Rauzy 2011 — De Looze C., Rauzy S. Measuring speakers' similarity in speech by means of prosodic cues: Methods and potential. *Proc. of Interspeech 2011*. Cosi P., De Mori R., Di Fabbrizio G., Pieraccini R. (eds.). International Speech Communication Association, 2011, 1393–1396. https://www.isca-speech.org/archive/interspeech_2011/i11_1393.html.
- De Looze et al. 2014 — De Looze C., Scherer S., Vaughan B., Campbell N. Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction. *Speech Communication*, 2014, 58: 11–34.
- Delvaux, Soquet 2007 — Delvaux V., Soquet A. The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica*, 2007, 64(2–3): 145–173.
- Edlund et al. 2009 — Edlund J., Heldner M., Hirschberg J. Pause and gap length in face-to-face interaction. *Proc. of Interspeech 2009*. Uther M., Moore R., Cox S. (eds.). International Speech Communication Association, 2009, 2779–2782. https://www.isca-speech.org/archive/interspeech_2009/i09_2779.html.
- Edlund 2011 — Edlund J. *In search for the conversational homunculus: Serving to understand spoken human face-to-face interaction*. Doctoral diss., Stockholm, 2011.
- Eyben et al. 2016 — Eyben F., Scherer K. R., Schuller B. W., Sundberg J., André E., Busso C., Devillers L. Y., Epps J., Laukka P., Narayanan S. S., Truong K. P. The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for voice research and affective computing. *IEEE Transactions on Affective Computing*, 2016, 7(2): 190–202.

- Fowler 1986 — Fowler C. A. An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 1986, 14(1): 3–28.
- Fowler et al. 2003 — Fowler C., Brown J., Sabadini L., Weihing J. Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language*, 2003, 49: 396–413.
- Giles 1973 — Giles H. Accent mobility: A model and some data. *Anthropological Linguistics*, 1973, 15(2): 87–105.
- Giles et al. 1973 — Giles H., Taylor D. M., Bourhis R. Towards a theory of interpersonal accommodation through language: Some Canadian data. *Language in Society*, 1973, 2: 177–192.
- Giles et al. 1977 — Giles H., Bourhis R. Y., Taylor D. M. Towards a theory of language in ethnic group relations. *Language, ethnicity and intergroup relations*. Giles H. (ed.). London: Academic Press, 1977, 307–348.
- Giles et al. 1991 — Giles H., Coupland N., Coupland J. Accommodation theory: Communication, context, and consequence. *Contexts of accommodation: Developments in applied sociolinguistics*. Giles H., Coupland J., Coupland N. (eds.). Cambridge: Cambridge Univ. Press, 1991, 1–68.
- Goldinger 1998 — Goldinger S. D. Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 1998, 105(2): 251–279.
- Gravano 2009 — Gravano A. *Turn-taking and affirmative cue words in task-oriented dialogue*. Ph.D. diss., Columbia Univ., 2009.
- Gregory, Webster 1996 — Gregory S. W. Jr., Webster S. A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of Personality and Social Psychology*, 1996, 70(6): 1231–1240.
- Haywood et al. 2005 — Haywood S. L., Pickering M. J., Branigan H. P. Do speakers avoid ambiguities during dialogue? *Psychological Science*, 2005, 16(5): 362–366.
- Heldner et al. 2010 — Heldner M., Edlund J., Hirschberg J. Pitch similarity in the vicinity of backchannels. *Proc. of Interspeech 2010*. Kobayashi T., Hirose K., Nakamura S. (eds.). International Speech Communication Association, 2010, 3054–3057.
- Hess, Blairy 2001 — Hess U., Blairy S. Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, 2001, 40(2): 129–141.
- Ireland et al. 2011 — Ireland M. E., Slatcher R. B., Eastwick P. W., Scissors L. E., Finkel E. J., Pennebaker J. W. Language style matching predicts relationship initiation and stability. *Psychological Science*, 2011, 22(1): 39–44.
- Karpiński et al. 2014 — Karpiński M., Klessa K., Czoska A. Local and global convergence in the temporal domain in Polish task-oriented dialogue. *Proc. of the 7th International Conf. on Speech Prosody*. Campbell N., Gibbon D., Hirst D. (eds.). International Speech Communication Association, 2014, 743–747. https://www.isca-speech.org/archive/SpeechProsody_2014/abstracts/141.html.
- Kawahara et al. 2015 — Kawahara T., Yamaguchi T., Uesato M., Yoshino K., Takanashi K. Synchrony in prosodic and linguistic features between backchannels and preceding utterances in attentive listening. *Proc. of Asia-Pacific Signal and Information Processing Association Annual Summit and Conf.* Lam K., Meng H., Au O. (eds.). Asia-Pacific Signal and Information Processing Association, 2015, 392–395.
- Kousidis et al. 2009 — Kousidis S., Dorran D., McDonnell C., Coyle E. Time series analysis of acoustic feature convergence in human dialogues. *Speech and Computer*. Karpov A. (ed.). St. Petersburg: St. Petersburg Institute for Informatics and Automation, 2009, 91–96.
- Lee et al. 2010 — Lee C.-C., Black M., Katsamanis A., Lammert A., Baucom B., Christensen A., Georgiou P., Narayanan S. Quantification of prosodic entrainment in affective spontaneous spoken interactions of married couples. *Proc. of Interspeech 2010*. Kobayashi T., Hirose K., Nakamura S. (eds.). International Speech Communication Association, 2010, 793–796.
- Levitan 2014 — Levitan R. *Acoustic-prosodic entrainment in human-human and human-computer dialogue*. Ph.D. diss., Columbia Univ., 2014.
- Levitan, Hirschberg 2011 — Levitan R., Hirschberg J. Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. *Proc. of Interspeech 2011*. Cosi P., De Mori R., Di Fabbriozio G., Pieraccini R. (eds.). International Speech Communication Association, 2011, 3081–3084.
- Levitan et al. 2015a — Levitan R., Beňuš Š., Gravano A., Hirschberg J. Acoustic-prosodic entrainment in Slovak, Spanish, English and Chinese: A cross-linguistic comparison. *Proc. of the SIGDIAL 2015 Conf., The 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. Koller A., Skantze G., Araki M., Rose C. P. (eds.). The Association for Computational Linguistics, 2015, 325–334.

- Levitan et al. 2015b — Levitan R., Beňuš Š., Gravano A., Hirschberg J. Entrainment and turn-taking in human-human dialogue. *Proc. of AAAI 2015 Spring Symposium on Turn-taking and Coordination in Human-Machine Interaction*. Andrist S., Bohus D., Horvitz E., Mutlu B., Schlangen D. (eds.). Palo Alto: The AAAI Press, 2015, 44–51.
- Levitan et al. 2016 — Levitan R., Beňuš Š., Gálvez R. H., Gravano A., Savoretti F., Trnka M., Weise A., Hirschberg J. Implementing acoustic-prosodic entrainment in a conversational avatar. *Proc. of Interspeech 2016*. Morgn N. (ed.). International Speech Communication Association, 2016, 1166–1170. https://www.isca-speech.org/archive/Interspeech_2016/abstracts/0985.html.
- Lewandowski, Jilka 2019 — Lewandowski N., Jilka M. Phonetic convergence, language talent, personality and attention. *Frontiers in Communication*, 2019, 4: Article 18.
- Lubold et al. 2015 — Lubold N., Pon-Barry H., Walker E. Naturalness and rapport in a pitch adaptive learning companion. *Proc. of 2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU 2015)*. Institute of Electrical and Electronics Engineers, 2015, 103–110. <https://ieeexplore.ieee.org/document/7404781>.
- Matarazzo, Wiens 1967 — Matarazzo J. D., Wiens A. N. Interviewer influence on durations of interviewee silence. *Journal of Experimental Research in Personality*, 1967, 2(1): 56–69.
- Maurer, Tindall 1983 — Maurer R. E., Tindall J. H. Effect of postural congruence on client's perception of counselor empathy. *Journal of Counseling Psychology*, 1983, 30(2): 158–163.
- Michalsky, Schoormann 2017 — Michalsky J., Schoormann H. Pitch convergence as an effect of perceived attractiveness and likability. *Proc. of Interspeech 2017*. Lacerda F. (ed.). International Speech Communication Association, 2017, 2253–2256. https://www.isca-speech.org/archive/Interspeech_2017/abstracts/1520.html.
- Namy et al. 2002 — Namy L., Nygaard L., Sauerteig D. Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology*, 2002, 21(4): 422–432.
- Nass, Lee 2001 — Nass C., Lee K. Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology Applied*, 2001, 7(3): 171–181.
- Natale 1975 — Natale M. Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, 1975, 32(5): 790–804.
- Nenkova et al. 2008 — Nenkova A., Gravano A., Hirschberg J. High frequency word entrainment in spoken dialogue. *Proc. of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short papers*. Moore J. D., Teufel S., Allan J., Furui S. (eds.). Association for Computational Linguistics, 2008, 169–172. <https://www.aclweb.org/anthology/P08-2043/>.
- Niederhoffer, Pennebaker 2002 — Niederhoffer K. G., Pennebaker J. W. Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, 2002, 21(4): 337–360.
- Oviatt et al. 2004 — Oviatt S., Darves C., Coulston R. Toward adaptive conversational interfaces: Modeling speech convergence with animated personas. *ACM Transactions on Computer-Human Interaction*, 2004, 11(3): 300–328.
- Pardo 2006 — Pardo J. S. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Association of America*, 2006, 119(4): 2382–2393.
- Pardo et al. 2013a — Pardo J. S., Cajori Jay I., Hoshino R., Hasbun S. M., Sowemimo-Coker C., Krauss R. M. The influence of role-switching on phonetic convergence in conversation. *Discourse Processes*, 2013, 50: 276–300.
- Pardo et al. 2013b — Pardo J. S., Jordan K., Mallari R., Scanlon C., Lewandowski E. Phonetic convergence in shadowed speech: The relation between acoustic and perceptual measures. *Journal of Memory & Language*, 2013, 69: 183–195.
- Pardo et al. 2017 — Pardo J., Urmanche A., Wilman Sh., Wiener J. Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics*, 2017, 79: 637–659.
- Pardo et al. 2019 — Pardo J. S., Urmanche A., Gash H., Wiener J., Mason N., Wilman S., Francis K., Decker A. The Montclair Map Task: Balance, efficacy, and efficiency in conversational interaction. *Language and Speech*, 2019, 62(2): 378–398.
- Pérez et al. 2016 — Pérez J., Gálvez R., Gravano A. Disentrainment may be a positive thing: A novel measure of unsigned acoustic-prosodic synchrony, and its relation to speaker engagement. *Proc. of Interspeech 2016*. Morgan N. (ed.). International Speech Communication Association, 2016, 1270–1274. https://www.isca-speech.org/archive/Interspeech_2016/abstracts/0587.html.
- Pickering, Garrod 2004 — Pickering M. J., Garrod S., Toward a mechanistic psychology of dialogue. *The Behavioral and Brain Sciences*, 2004, 27(2): 169–190 (Discussion 190–226).

- Ray, Webb 1966 — Ray M. L., Webb E. J. Speech duration effects in the Kennedy news conferences. *Science*, 1966, 153: 899–901.
- Reichel 2014 — Reichel U. Linking bottom-up intonation stylization to discourse structure. *Computer, Speech, and Language*, 2014, 28: 1340–1365.
- Reichel, Cole 2016 — Reichel U., Cole J. Entrainment analysis of categorical intonation representations. *Proc. of Phonetik und Phonologie im deutschsprachigen Raum*. Draxler C., Kleber F. (eds.). Munich: Ludwig-Maximilians-Universität München, 2016, 165–168.
- Reichel et al. 2018 — Reichel U., Beňuš Š., Mady K. Entrainment profiles: Comparison by gender, role, and feature set. *Speech Communication*, 2018, 100: 46–57.
- Reitter et al. 2006 — Reitter D., Moore J. D., Keller F. Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. *Proc. of the 28th Annual Conf. of the Cognitive Science Society*. Sun R., Miyake N. (eds.). Vancouver: Cognitive Science Society, 2006, 717–722.
- Savino et al. 2016 — Savino M., Lapertosa L., Caffò A., Refice M. Exploring prosodic convergence in Italian game dialogues. *Proc. of the 7th Tutorial & Research Workshop on Experimental Linguistics (Exling 2016)*. Botinis A. (ed.). International Speech Communication Association, 2016, 151–154.
- Schweitzer, Lewandowski 2013 — Schweitzer A., Lewandowski N. Convergence of articulation rate in spontaneous speech. *Proc. of Interspeech 2013*. Bimbot F., Cerisara C., Fougeron C., Gravier G., Lamel L., Pellegrino F., Perrier P. (eds.). International Speech Communication Association, 2013, 525–529. https://www.isca-speech.org/archive/interspeech_2013/i13_0525.html.
- Shockley et al. 2004 — Shockley K., Sabadini L., Fowler C. A. Imitation in shadowing words. *Perception & Psychophysics*, 2004, 66: 422–429.
- Suzuki, Katagiri 2007 — Suzuki N., Katagiri Y. Prosodic alignment in human-computer interaction. *Connection Science*, 2007, 19: 131–141.
- Székely et al. 2015 — Székely É., Keane M., Carson-Berndsen J. The effect of soft, modal and loud voice levels on entrainment in noisy conditions. *Proc. of Interspeech 2015*. Möller S., Ney H. (eds.). International Speech Communication Association, 2015, 150–154. https://www.isca-speech.org/archive/interspeech_2015/i15_0150.html.
- Trancoso et al. 1998 — Trancoso I., do Céu Viana M., Duarte I., Matos G. Corpus de diálogo CORAL. *Proc. of Computational Processing of the Portuguese Language*. Porto Alegre, 1998, 99–108.
- Vaughan 2011 — Vaughan B. Prosodic synchrony in co-operative task-based dialogues: A measure of agreement and disagreement. *Proc. of Interspeech 2011*. Cosi P., De Mori R., Di Fabrizio G., Pieraccini R. (eds.). International Speech Communication Association, 2011, 1865–1868. https://www.isca-speech.org/archive/interspeech_2011/i11_1865.html.
- Ward, Litman 2007 — Ward A., Litman D. Automatically measuring lexical and acoustic/prosodic convergence in tutorial dialog corpora. *Proc. of the Workshop on Speech and Language Technology in Education (SLaTE2007)*. Farmington: Carnegie Mellon Univ., 2007, 57–60. https://www.isca-speech.org/archive_open/slate_2007/sle7_057.html.
- Ward, Nakagawa 2002 — Ward N., Nakagawa S. Automatic user-adaptive speaking rate selection for information delivery. *Proc. of International Conf. on Spoken Language Processing (ICSLP 2002)*. Hansen J., Pellom B. (eds.). Denver: Casual Prod. Ltd., 2002, 549–552.
- Weidman et al. 2016 — Weidman S., Breen M., Haydon K. Prosodic speech entrainment in romantic relationships. *Proc. of the 8th International Conf. on Speech Prosody*. Barnes J., Brugos A., Shattuck-Hufnagel S., Veilleux N. (eds.). International Speech Communication Association, 2016, 508–512.
- Weise et al. 2019 — Weise A., Levitan S. I., Hirschberg J., Levitan R. Individual differences in acoustic-prosodic entrainment in spoken dialogue. *Speech Communication*, 2019, 115: 78–87.
- Welkowitz et al. 1972 — Welkowitz J., Feldstein S., Finkelstein M., Ayelsworth L. Changes in vocal intensity as a function of interspeaker influence. *Perceptual and Motor Skills*, 1972, 35: 715–718.
- Xia et al. 2014 — Xia Z., Levitan R., Hirschberg J. Prosodic entrainment in Mandarin and English: A cross-linguistic comparison. *Proc. of the 7th International Conf. on Speech Prosody*. Campbell N., Gibbon D., Hirst D. (eds.). International Speech Communication Association, 2014, 65–69.
- Zellers 2016 — Zellers M. Prosodic convergence with spoken stimuli in laboratory data. *Proc. of Interspeech 2016*. Morgan N. (ed.). International Speech Communication Association, 2016, 1021–1025.
- Zoltan-Ford 1991 — Zoltan-Ford E. How to get people to say and type what computers can understand. *International Journal of Man-Machine Studies*, 1991, 34(4): 527–547.