



# Моделирование речевой просодии: временной компонент выделительного акцента в английском языке

*Филясова Ю.А.*

В предлагаемой статье рассматривается реализация акцентного выделения (АВ) в английском языке на участках гласных в составе односложных слов, выделенных из кратких диалогов, в чтении одного диктора. Автор обнаруживает различия в длительности гласных в зависимости от позиции акцентированного слова во фразе и глухости/звонкости конечного согласного слова и делает вывод о том, что выделительный акцент в английском языке реализуется в условиях взаимодействия просодических и сегментных факторов. С целью возможного практического применения полученных данных выделен процент прироста длительности гласных, ответственный за реализацию АВ в заданных условиях.

## Введение

Речевые технологии как одно из направлений информационных технологий занимается проблемами общения человека с компьютером (или человека с человеком посредством компьютера) на основе использования языка в его звуковой форме. Речевые технологии используются в создании систем автоматического синтеза и распознавания речи; при построении современных средств речевого общения со сложными техническими устройствами, занимающими всё более значительное место в разных сферах жизни общества [1], в частности, при создании человеко-машинных интерфейсов с устным вводом/выводом информации, при организации информационно-справочной службы, в обучении иностранному языку (автоматические фонетические тренажёры) и т.д.

Современные программные и технические средства позволяют моделировать естественные процессы порождения и восприятия речи. При этом необходимо помнить, что эффективность моделирования речевого сигнала и

решения других прикладных задач определяется полнотой использования фонетических сведений, полученных при изучении свойств естественной речи. Особое значение для речевых технологий представляют сведения о просодической организации устной речи, которые ответственны за естественность звучания синтезированной речи. Так, например, формирование фразовых интонационно-просодических показателей (типов интонации, особых подчёркиваний слов и т.д.) является одной из основных задач на этапе лингвистической обработки в системах «Текст — Речь» [2]. Изучение просодических явлений осложняется их внутренней взаимозависимостью на различных уровнях преобразования (лексическом, фонетическом, прагматическом). Каждый из просодических параметров должен извлекаться в результате довольно сложного анализа, который сам по себе варьируется в зависимости от степени продвинутой инструментария, речевых технологий и т.д. [3].

В данной статье приводятся некоторые данные исследования длительности гласных в условиях акцентного выделения<sup>1</sup>. АВ является языковым средством подчёркивания особо важных элементов высказывания. Местоположение его в высказывании не фиксировано, поскольку обусловлено конкретными коммуникативными задачами.

Длительность, по данным современных исследований, является важной супraseгментной характеристикой выделения основных и фоновых (второстепенных) смысловых фрагментов текста. Сегментная длительность служит одним из признаков эмфатического или противопоставительного ударения [5]. Поскольку во многих языках длительность выполняет функцию различения фонем, долгое время её не рассматривали в качестве одного из главных просодических параметров. Одним из первых, кто начал рассматривать длительность как один из наиболее информативных компонентов просодии, был Д. Фрай [6]. Он делает вывод о том, что увеличение длительности в английском языке — более эффективное средство выделения, по сравнению с увеличением интенсивности.

В изучении акустических параметров АВ в английском языке длительность приобретает особое значение, поскольку определённым образом участвует в сегментной организации речи. Ещё в XIX веке в работах по фонетике учёные писали о сокращении длительности гласных в контексте перед глухими согласными [7]. Современные исследователи отмечают регулярный характер данного явления в английском языке [8; 9, с. 58].

Существует также универсальная фонетическая закономерность — увеличение длительности сегментов к концу коммуникативной единицы, — которая характеризует любое устное высказывание и не зависит от сегментного состава единиц, входящих в это высказывание. Так, например, Д. Клатт определил, что в английском языке длительность гласных в области перед паузой увеличивается на 30% [10].

Кроме того, особенностью английского языка является строгая ритмическая организация устной речи («stress-timing»). Одним из первых её отметил Д. Аберкромби [11]. Ударные слоги реализуются через примерно равные промежутки времени, независимо от количества безударных слогов между ними. Длительность при этом выступает в качестве средства для оформления равных ритмических отрезков высказывания.

С учётом указанных функций темпорального компонента в английском языке, сведения об изменениях длительности гласных английского языка в условиях акцентного выделения могут внести значительный вклад в лингвистические основания речевых технологий.

<sup>1</sup> Мы используем термин «акцентное выделение» (далее — АВ) вслед за Т.М. Николаевой, см. [4].



## Эксперимент

В данной статье описываются результаты исследования, **задача** которого состояла в том, чтобы выявить особенности реализации длительности гласных при АВ в условиях влияния таких фонетических факторов, как: 1) сегментный (правый) консонантный контекст; 2) фразовая позиция.

**Материалом** исследования послужила запись микродиалогов в чтении носителей британского варианта английского языка, которая выполнялась в Лаборатории экспериментальной фонетики (ЛЭФ) им. акад. Л.В. Щербы СПбГУ. Для анализа был отобран материал одного диктора. Исследованию подлежали гласные в словах структуры CVC<sup>2</sup> в разных фразовых позициях (табл. 1). Длительность гласных в неакцентных позициях выступает в качестве объекта для сравнения длительности гласных в акцентных позициях. Нейтральной позицией, т.н. точкой отсчёта, стала «IA-»<sup>3</sup> (на рис. 3 отмечена овалом). Для обработки материала использовались программы Wave Assistant v/2.00 (Центр Речевых Технологий (ЦРТ), Санкт-Петербург, 1998), Excel, Statistica v/5.00. Слова условно делятся на две группы: со звонким (V+) и глухим (V-) конечным согласным, например, bag — back, cad — cat. Количество реализаций гласных в каждой заданной позиции насчитывает 30 ед., как в контексте перед глухими, так и в контексте перед звонкими согласными. Общее количество исследованных гласных составляет 360 ед.<sup>4</sup>

Таблица 1

Фразовые позиции для слова

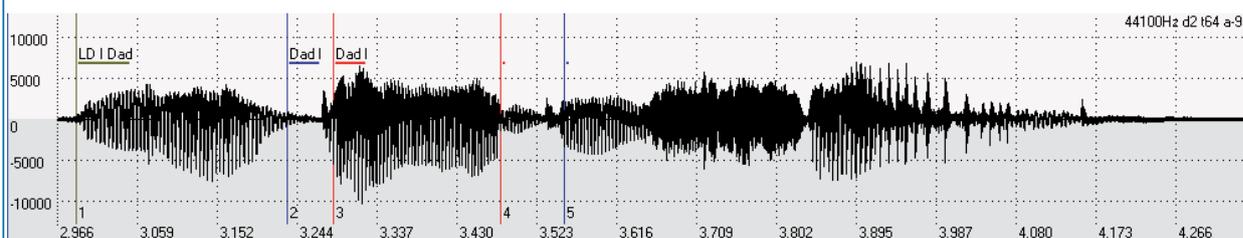
№	Условное обозначение позиции	Позиция слова во фразе относительно выделенности	Образцы микродиалогов на примере слова «bag»
1	«I» (Initial)	Слово под ударением	<i>The «bag» was found.</i>
2	«IA-» (Initial Accented)	Слово в позиции после выделенного слова	A: <i>Whose bag was found? Was your bag found?</i> B: <i>No, <b>your</b> «bag» was found.</i>
3	«IA+» (Initial Accented)	Слово с выделительным ударением	A: <i>What was found? A sack?</i> B: <i>No, a «<b>bag</b>» was found.</i>
4	«F» (Final)	Слово под синтагматическим ударением	<i>This is a «bag».</i>
5	«FA-» (Final Accented)	Слово в позиции после выделенного слова	A: <i>What a small sack you bought!</i> B: <i>But this is not a sack, this is a bag.</i> A: <i>Since I took it for a sack, it is a <b>very big</b> «bag».</i>
6	«FA+» (Final Accented)	Слово с выделительным ударением	A: <i>What a small sack you bought!</i> B: <i>But this is not a sack, this is a «<b>bag</b>».</i>

<sup>2</sup> CVC: Consonant — Vowel — Consonant (Согласный — Гласный — Согласный).

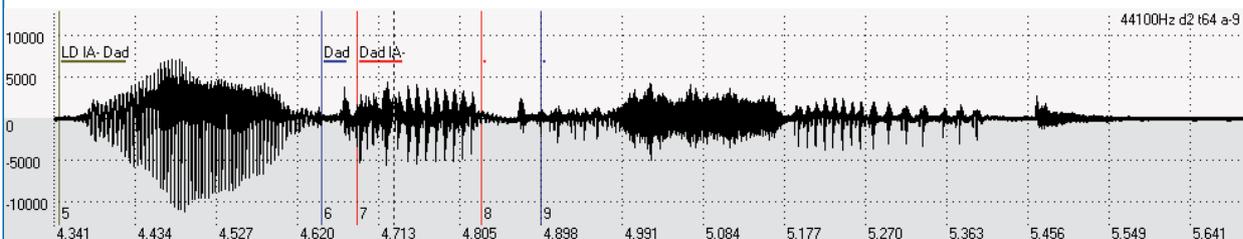
<sup>3</sup> Мы считаем, что гласный в безударной позиции в начале фразы менее всего подвержен влиянию различных просодических факторов.

<sup>4</sup> В данной статье представлены результаты исследования длительности гласных в составе утвердительных фраз.

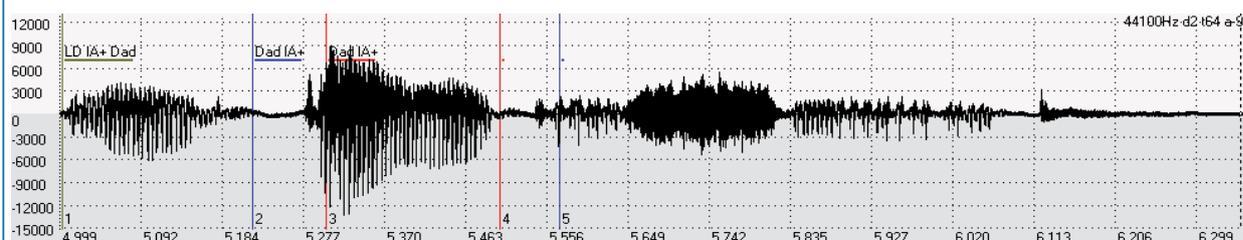
На рис. 1 и 2 представлены примеры фраз, содержащие отобранные для исследования слова. Каждый рисунок демонстрирует слово в одной из заданных фразовых позиций. Для примеров выбраны два слова, контрастные по сегментному составу: «Dad» vs. «tar». Оба слова имеют одинаковый центральный элемент — открытый гласный переднего ряда /æ/, но в одном случае перед глухим, а в другом — перед звонким согласным. Изменения длительности гласного можно проследить не только от начала к концу фразы (фразово-позиционные изменения), но и в одинаковых фразовых позициях перед разными согласными (сегментные различия).



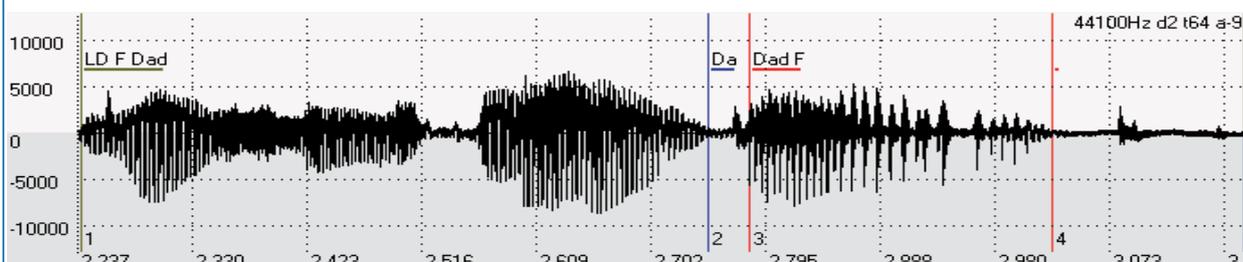
A: (Позиция «I») My «Dad» is sad<sup>5</sup>. (/æ/=194 мс)



B: (Позиция «IA-») Your «Dad» is sad. (/æ/=143 мс)



C: (Позиция «IA+») My «Dad» is sad. (/æ/=199 мс)



D: (Позиция «F») This is my «Dad». (/æ/=245 мс)

<sup>5</sup> Слова для исследования взяты в кавычки. «Жирным» шрифтом отмечены слова с АВ. Слова в кавычках «жирным» шрифтом были отобраны для исследования и произнесены в составе диалогов с АВ.

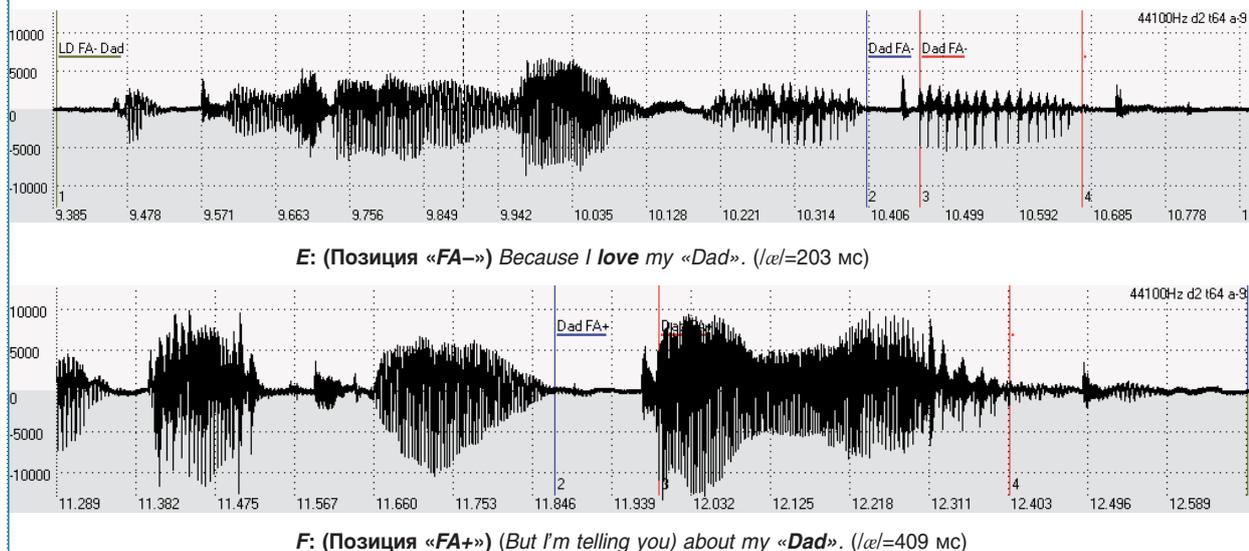


Рис. 1. Длительность гласного / $\alpha$ / в заданных фразовых позициях в составе слова «Dad»<sup>6</sup>

На примере слова «Dad» видно, что длительность гласного в конечных позициях («F», «FA-», «FA+») больше, чем в начальных («I», «I-», «I+»). Примеры демонстрируют проявление общеизвестной фонетической закономерности увеличения длительности к концу высказывания. Наибольшая длительность наблюдается в конечной акцентной позиции «FA+» (409 мс). Следует отдельно подчеркнуть, что длительность гласных в словах с АВ в конце фразы на 184 мс больше, чем под нейтральным синтагматическим ударением без АВ (409 мс в позиции «FA+» vs. 225 мс в позиции «F»).

В начальной акцентной позиции «I+» длительность гласного / $\alpha$ / не намного больше, чем в позиции без АВ «I» (199 мс vs. 194 мс), и является наибольшей только среди начальных позиций. Таким образом, длительность гласных при АВ зависит от позиции во фразе. Эта зависимость проявляется и на длительности гласных в сегментном контексте перед глухим согласным (см. рис. 2).

Сравнение длительности гласных в одинаковых позициях, но с разным правым консонантным контекстом (рис. 1 vs. рис. 2) показывает, что в слове «tap» длительность гласных перед глухим согласным во всех заданных позициях значительно меньше, чем в слове «Dad». Позиции с АВ не являются исключением. Так, например, в позиции «I+»: 199 мс в слове «Dad» vs. 90 мс в слове «tap». В позиции «FA+»: 409 мс в слове «Dad» vs. 116 мс в слове «tap».

<sup>6</sup> Для наглядного представления фраз были заданы параметры, одинаковые для всех примеров: частота дискретизации 44100 Гц, растяжение по оси ОХ — 64, растяжение по оси ОУ — 9. Синими метками обозначены слова, красными метками — гласные, жёлтыми метками указаны границы фраз. При этом частота дискретизации понимается как частота взятия отсчётов непрерывного во времени сигнала при его дискретизации аналогоцифровым преобразователем [12].

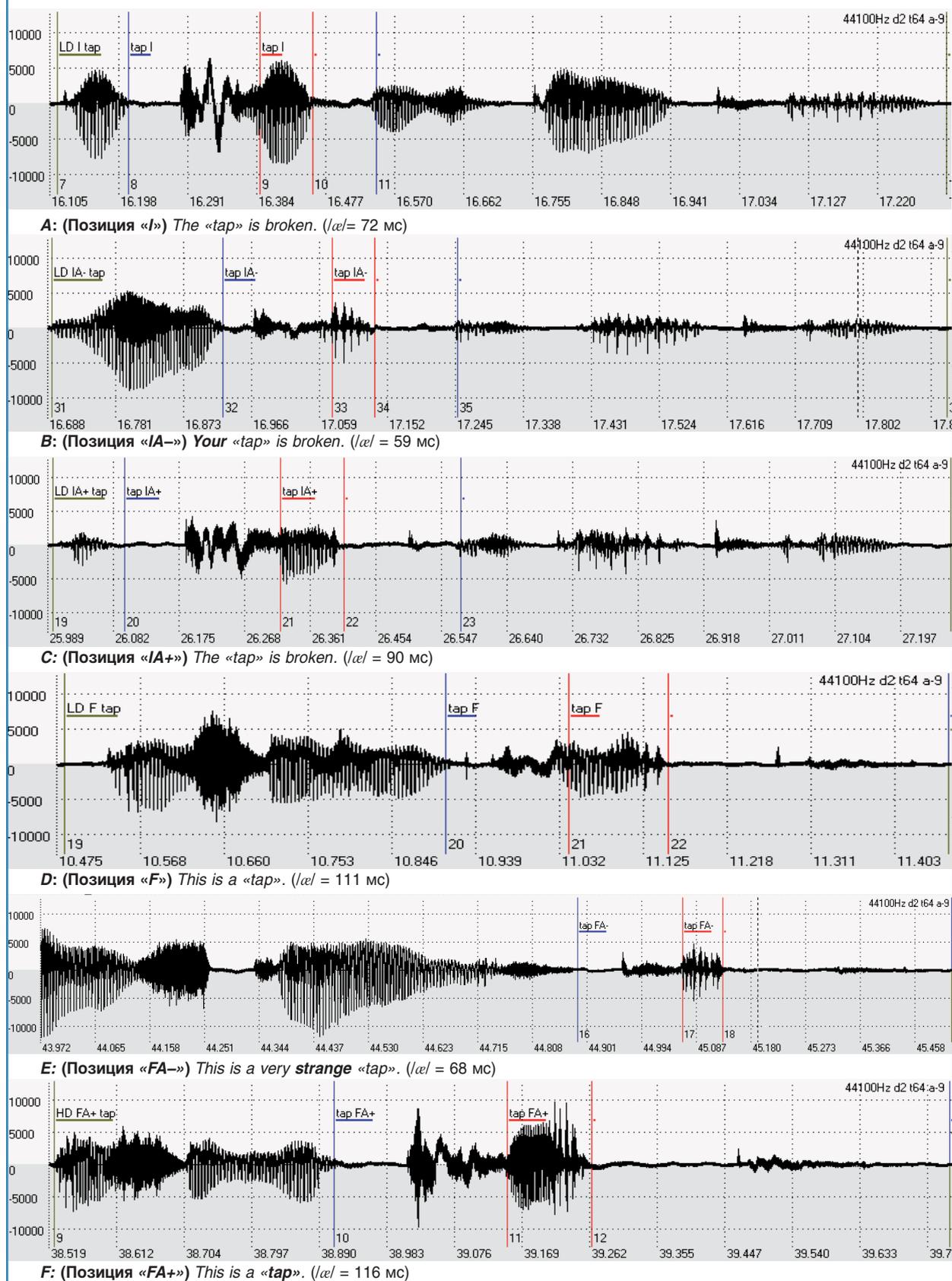


Рис. 2. Длительность гласного /æ/ в заданных фразовых позициях в составе слова «tap»

Поведение длительности гласных на примере указанных выше двух слов вполне соотносится со средними показателями длительности, полученными на материале 360 гласных.

Рассмотрим фразово-позиционные изменения длительности гласных в контексте перед звонкими и глухими согласными.

При больших значениях длительности гласных перед звонкими<sup>7</sup> согласными увеличение показателей длительности гласных наблюдается от начала фразы к её концу, как в контексте перед звонкими, так и в контексте перед глухими согласными (см. рис. 3). Так, перед звонкими согласными средняя длительность гласных изменяется от 140 мс (в начальной безударной позиции «IA-») до 330 мс (в конечной акцентной позиции «FA+»); перед глухими — от 90 мс (в начальных ударной «I» и безударной «IA-» позициях) до 170 мс (в конечной акцентной позиции «FA+»).

Длительность гласных в начальной акцентной позиции «IA+» в обоих сегментных контекстах меньше длительности гласных в конечной акцентной позиции «FA+» (на 100 мс в контексте перед звонкими согласными и на 50 мс — перед глухими).

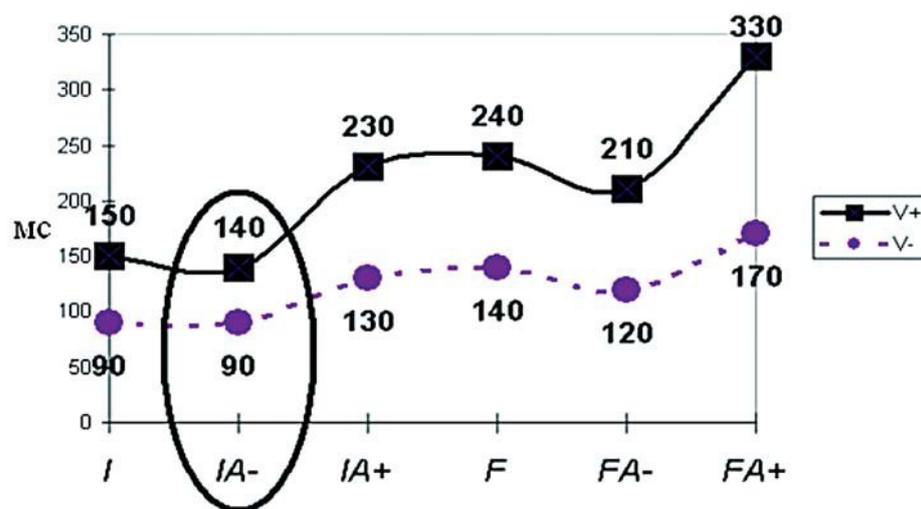


Рис. 3. Средняя длительность гласных в заданных фразовых позициях в контексте перед звонкими и глухими согласными, в миллисекундах («V+» — звонкий правый консонантный контекст, «V-» — глухой правый консонантный контекст)

Увеличение длительности гласных в заданных фразовых позициях относительно длительности гласных в нейтральной позиции «IA-» статистически достоверно как в контексте перед звонкими, так и в контексте перед глухими согласными. Исключение составляет начальная ударная позиция

<sup>7</sup> О большей длительности гласных перед звонкими согласными в английском языке см., например, вступительную статью в [13].

без АВ «/» (для V-), т.к. длительность гласных в сравниваемых позициях оказалась одинаковой (см. табл. 2, где статистически достоверные различия отмечены подчеркнутым курсивом).

Таблица 2

Тип сегментного контекста	V+	V+	V+	V+	V+	V-	V-	V-	V-	V-
Позиции в синтагме	I	IA+	F	FA-	FA+	I	IA+	F	FA-	FA+
t-критерий	<u>2.63</u>	<u>10.06</u>	<u>11.65</u>	<u>9.72</u>	<u>11.03</u>	0,79	<u>9.68</u>	<u>7.58</u>	<u>3.15</u>	<u>7.59</u>
p<0,05	<u>0.01</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	0,43	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>

t-критерий — критерий значимости изменения длительности гласных,  
p — вероятность появления случайной величины вне допустимого интервала варьирования значений в заданной выборке.

Относительное увеличение длительности гласных, в зависимости от позиции во фразе, изменяется в диапазоне от 7% до 136% в контексте перед звонкими согласными и от 0% до 89% — перед глухими (см. рис. 4). Как видно, в контексте перед звонкими согласными оно больше: в начальной акцентной позиции «IA+» — на 20%; в конечной позиции синтагматического ударения «F» — на 15%; в конечной безударной позиции «FA-» — на 17%; в конечной акцентной позиции «FA+» — на 47%.

Длительность гласных в начальной акцентной позиции «IA+» оказалась не только меньше длительности гласных в конечной акцентной позиции «FA+» (на 72% перед звонкими согласными и на 45% — перед глухими), но и меньше длительности гласных под нейтральным синтагматическим ударением «F» (на 7% в контексте перед звонкими согласными и на 14% — перед глухими). Это подчёркивает значимость конечной фразовой позиции для длительности отдельных элементов.

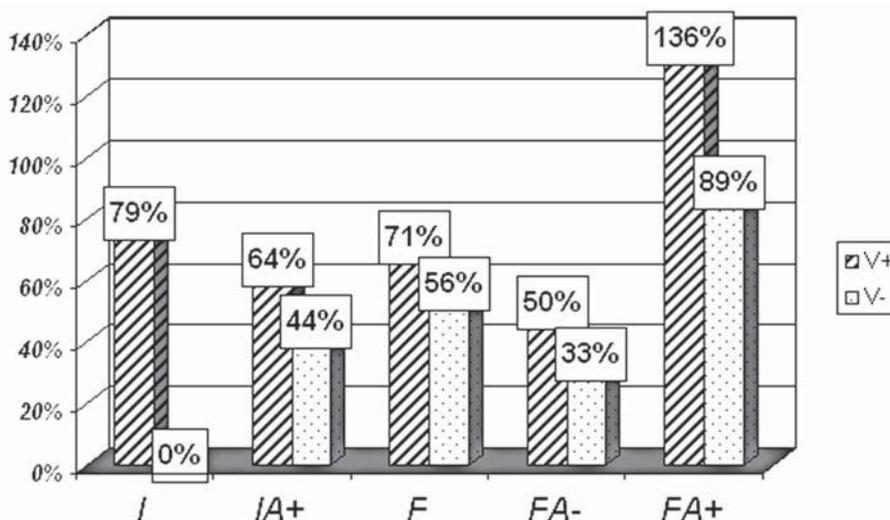


Рис. 4. Увеличение длительности гласных в заданных фразовых позициях в контексте перед звонкими и глухими согласными относительно длительности гласных в нейтральной позиции «IA-», в процентах



Максимальное увеличение длительности гласных в конечной акцентной позиции «FA+» можно объяснить действием «кумулятивного эффекта»<sup>8</sup>. Как было отмечено выше, по данным некоторых исследователей, увеличение длительности гласных под синтагматическим ударением достигает 30% [8]. Материал, исследованный в данной работе на участках гласных, показывает, что при АВ в конце фразы к известным 30% добавляется еще 100% в контексте перед звонкими согласными и 60% — перед глухими согласными.

Более детальный анализ статистически достоверных различий в длительности гласных показывает, что диапазон стандартного отклонения и доверительного интервала соотносятся по-разному для гласных в разных фразовых позициях и в разных сегментных контекстах.

Так, в контексте перед звонкими согласными (V+) в конечной акцентной позиции статистически значимыми различия являются за счёт полного расхождения величин стандартного отклонения, в результате чего вероятность совпадения длительности гласных в данном случае равна нулю (рис. 5, B). В отношении начальной акцентной позиции «IA+» следует отметить, что существует вероятность совпадения значений с показателями длительности гласных из нейтральной позиции, учитывая общую границу размахов стандартного отклонения в двух выборках (рис. 5, A). Таким образом, большее приращение длительности гласных соотносится с меньшей степенью вероятности совпадения значений.

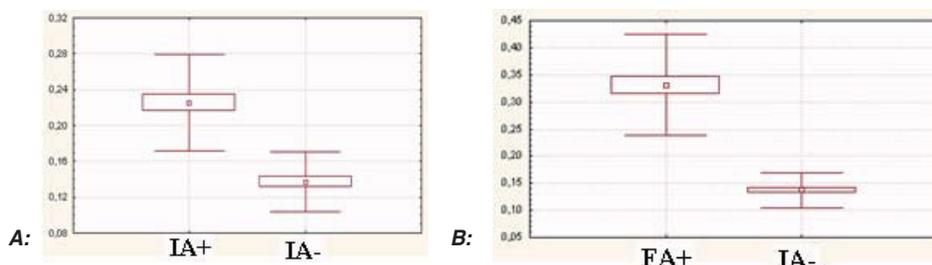


Рис. 5. Стандартное отклонение и доверительный интервал варьирования величин в заданных выборках

Разность в относительном приращении длительности гласных между акцентными позициями (72%) является статистически значимой,  $p=0,00$  (рис. 6). Допустимые интервалы варьирования значений длительности в рассматриваемых позициях разные. Диапазоны стандартного отклонения при этом частично совпадают, что оставляет некоторую долю вероятности совпадения значений из двух выборок.

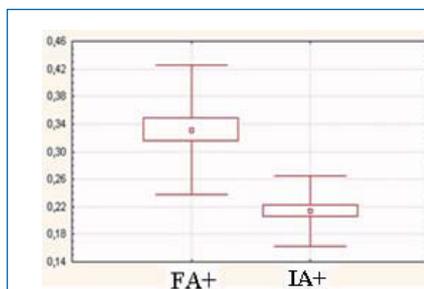


Рис. 6. Стандартное отклонение и доверительный интервал варьирования величин в заданных выборках

<sup>8</sup> О невозможности реализации «кумулятивного эффекта» см. в работе [14].

В контексте перед глухими согласными (V-) увеличение длительности гласных статистически достоверно в обеих акцентных позициях, благодаря разным значениям доверительных интервалов. Совпадение стандартных отклонений, однако, больше в начальной акцентной позиции «IA+» (рис. 7, А), чем в конечной «FA+» (рис. 7, В). В целом, совпадение стандартного отклонения по позициям больше, чем в контексте перед звонкими согласными (ср. рис. 7 с рис. 5), что можно объяснить меньшим приростом в относительной длительности гласных.

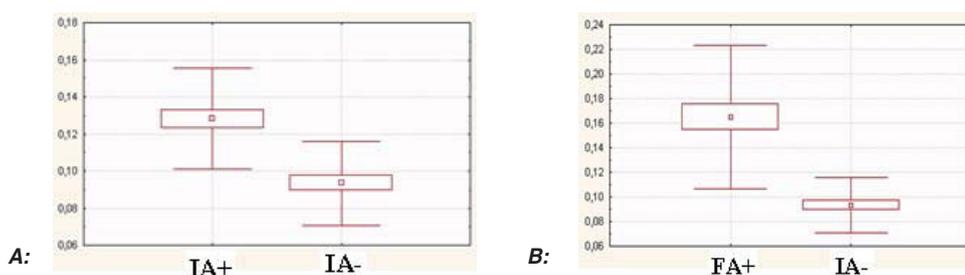


Рис. 7. Стандартное отклонение и доверительный интервал варьирования величин в заданных выборках

Разница в относительном увеличении длительности гласных между акцентными позициями «IA+» vs. «FA+» (45%) является статистически существенной ( $p=0,000214$ ) (рис. 8), так же как и в контексте перед звонкими согласными. При этом области стандартного отклонения в акцентных позициях значительно совпадают (ср. рис. 8 с рис. 6). Статистическая существенность различий обеспечивается разными допустимыми интервалами варьирования абсолютных величин.

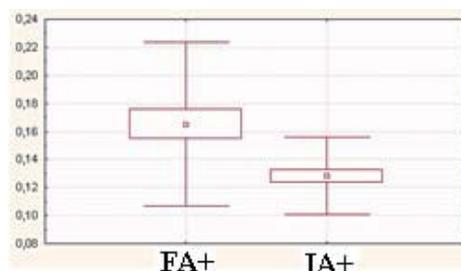


Рис. 8. Стандартное отклонение и доверительный интервал варьирования величин в заданных выборках

Рассмотрим сегментно-позиционные изменения длительности гласных в заданных фразовых позициях.

Разность в средней длительности гласных в контексте перед звонкими vs. глухими согласными варьируется от 50 до 160 мс (табл. 3). Наибольшая разность получена в позициях с АВ: «IA+» (100 мс) и «FA+» (160 мс) – и в позиции синтагматического ударения без АВ «F» (100 мс).

Таблица 3

Позиции в синтагме	I	IA-	IA+	F	FA-	FA+
(V+)-(V-), мс	60	50	100	100	90	160
t-критерий	<u>10,36</u>	<u>6,87</u>	<u>10,10</u>	<u>11,93</u>	<u>12,99</u>	<u>9,94</u>
p<0,05	<u>0,00</u>	<u>0,00</u>	<u>0,00</u>	<u>0,00</u>	<u>0,00</u>	<u>0,00</u>



Согласно t-критерию Стьюдента, полученная разница во всех случаях является статистически значимой. Диапазон стандартного отклонения показывает, что вероятность совпадения значений длительности гласных в сравниваемых выборках равна нулю (рис. 9). Следовательно, в акцентных позициях длительность гласных перед звонкими согласными всегда больше, чем перед глухими. Длительность гласных перед глухими согласными при появлении АВ не достигает значения длительности гласных перед звонкими согласными.

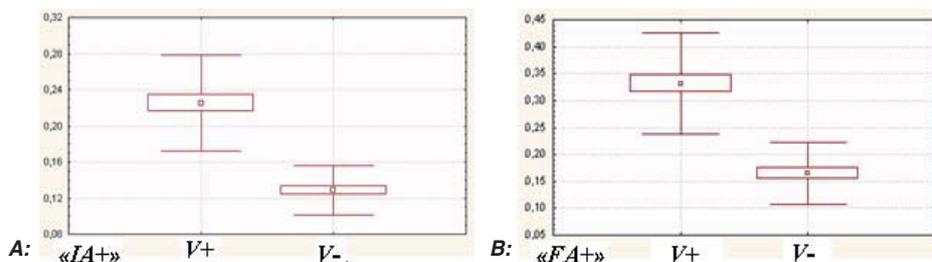


Рис. 9. Стандартное отклонение и доверительный интервал варьирования величин в заданных выборках

Относительное увеличение варьируется от 56% до 94%, в зависимости от позиции во фразе (рис. 10). При АВ разница в длительности гласных составляет 77% (в позиции «IA+») и 94% (в позиции «FA+»). Следовательно, длительность гласных в глухом сегментном контексте меньше длительности гласных в звонком сегментном контексте более чем на 50%<sup>9</sup>.

Полученные данные позволяют предположить, что больший прирост абсолютной длительности гласных обуславливает большую разницу в сегментной длительности гласных (перед звонкими vs. глухими согласными). Так, например, относительное фразово-позиционное увеличение длительности гласных в конечной акцентной позиции («FA+») больше, чем в начальной позиции с АВ («IA+»). Разница в сегментной длительности гласных в конечной акцентной позиции также оказалась больше, чем в начальной акцентной позиции (94% vs. 77%) (см. рис. 10).

## Выводы

Результаты исследования длительности гласных в условиях АВ позволяют сделать следующие выводы:

- относительное фразово-позиционное увеличение длительности гласных в АВ является значительным и статистически достоверным в контексте как перед звонкими, так и перед глухими согласными;
- в контексте перед звонкими согласными прирост длительности гласных больше, чем перед глухими, не только в абсолютных, но и в относительных показателях;
- в конечной акцентной позиции фразово-позиционное увеличение длительности гласных вдвое больше, чем в начальной, в контексте как перед

<sup>9</sup> Следует отметить, что сегментные различия в длительности гласных не выполняют функции различения гласных фонем в английском языке (ср.: «In most varieties of English, variations in lengths are completely allophonic» [9, с.225]).

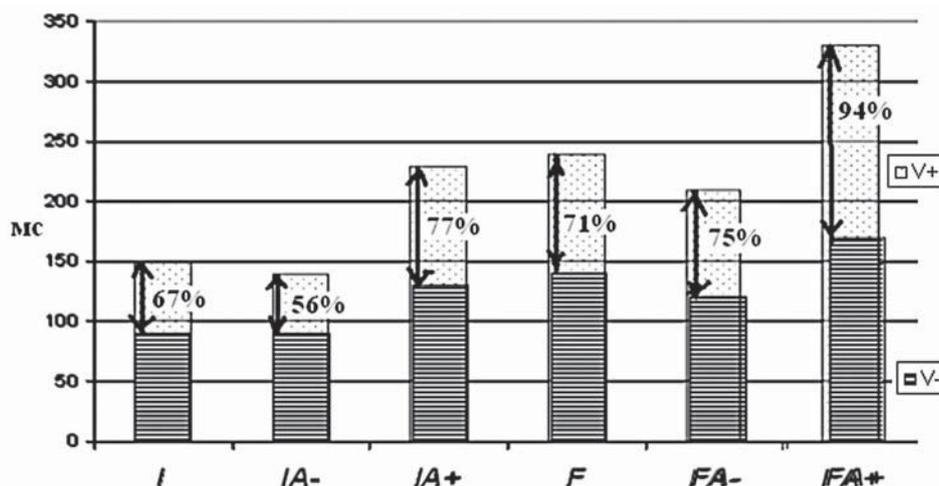


Рис.10. Увеличение длительности гласных в контексте перед звонкими согласными, по сравнению с длительностью гласных в контексте перед глухими согласными, в процентах

звонкими, так и перед глухими согласными; различия между начальной vs. конечной акцентными позициями статистически значимы;

— при появлении АВ в конечной фразовой позиции наблюдается действие «кумулятивного эффекта», когда к известным 30% [8] в позиции нейтрального синтагматического ударения может прибавиться ещё вплоть до 100% длительности гласных в контексте перед звонкими согласными и вплоть до 60% в контексте перед глухими согласными; таким образом, можно предположить, что именно эти дополнительные проценты ответственны за реализацию АВ в конечной фразовой позиции;

— в акцентных позициях сегментные различия по длительности становятся более яркими. Согласно полученным данным, в акцентных позициях длительность гласных перед звонкими согласными больше, чем перед глухими, на 77%–94%, а в неакцентных позициях — на 56–75%. Среди акцентных позиций более яркие сегментно-позиционные различия проявляются в конце фразы, где происходит наибольшее увеличение абсолютной длительности гласных как перед звонкими, так и перед глухими согласными.

Таким образом, при акцентном выделении вместе с увеличением длительности к концу фразы очень чётко проявляется аллофонное варьирование сегментно-позиционной длительности гласных. Полученные результаты свидетельствуют о совместном влиянии просодических и сегментных факторов на участках реализации АВ и вполне соответствуют выводам Л.А. Чистович и др.: «Реализация логических ударений производится путём таких преобразований артикуляторной программы произносимых предложений, которые не затрагивают её структурного существа, но могут быть описаны как введение некоторых коэффициентов, производящих закономерную деформацию элементов этой программы» [15].

## Литература

1. Скрелин П.А. Фонетические аспекты речевых технологий. Дисс. докт. филол. наук. СПб.: Изд-во СПбГУ, 1999. 78 с. С. 3–5.
2. Кодзасов С.В., Кривнова О.Ф. Общая фонетика. М., 2001. 592 с. С. 515–524.



3. Потапова Р.К. Язык, речь, личность. М.: Языки славянской культуры, 2006. 491 с. С. 11.
4. Николаева Т.М. Семантика акцентного выделения. М., 2004.
5. Klatt D. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence // *JASA*, 59, 1976, pp.1208–1221.
6. Fry D.B. (1955). Duration and intensity as physical correlates of linguistic stress // *JASA*, 27, 765–768.
7. Гордина М.В. История фонетических исследований (от античности до возникновения фонологической теории). СПб., 2006. 538 с. С. 306–335.
8. Gimson A.C. An introduction to the pronunciation of English. London, 1966. 294 p. С. 266.
9. Ladefoged P. A course in phonetics. San Diego, 1982. 30 p.
10. Klatt D.H. Vowel lengthening is syntactically determined in a connected discourse // *Journal of Phonetics* 3, 1975, pp.129–140.
11. Halliday M.A.K. Intonation and Grammar in British English. The Hague, 1967. 62 с. С. 15.
12. <http://ru.wikipedia.org/>
13. Jones D. An English Pronouncing Dictionary, 17<sup>th</sup> edn., P.Roach, J.Hartman and J.Setter (eds.). Cambridge: CUP, 2006.
14. Cooper W.E., Eady S.J., Mueller P.R.J. Acoustical aspects of contrastive stress in question-answer context // *JASA*, (1985), Vol. 77, No. 6, pp. 2142–2156.
15. Чистович Л.А. и др. Речь: артикуляция и восприятие. М.; Л., 1965.

**Филясова Юлия Анатольевна**

аспирант кафедры фонетики и методики преподавания иностранных языков СПбГУ,  
младший научный сотрудник Лаборатории экспериментальной фонетики им. Л.В. Щербы  
Института филологических исследований (ЛЭФ ИФИ).  
yuliyafill@gmail.com