

Синтез просодических моделей в языке дари

Иванов В.Б., доктор филологических наук, профессор



Речь на афганском дари имеет систему фонетических отличий от близкородственных языков (персидского и таджикского). В этой системе значительное место принадлежит просодическим моделям — другому типу словесного ударения. Если в персидском языке ударение чисто тоническое, а в таджикском по преимуществу тоническое, то в дари оно по преимуществу квантитативное, т.е. ударный слог, как правило, выделяется большей длительностью (по данным анализа ударный гласный удлиняется примерно на 10%, ударный слог — на 35%). Синтез речи позволил определить количественные границы этого противопоставления (даже 10-процентного удлинения слога достаточно для восприятия его ударным). Контур тона оказался незначим для ударения. Квантитативный компонент ударения может быть нейтрализован 25-процентным приростом тона и/или интенсивности в безударном слоге.

• Дари • иранские языки • персидский • таджикский • фарси • синтез • просодика • ударение.

Afghan Dari is characterized by a system of phonetic differences from closely related languages Persian and Tajik. In this system significant role is played by prosody and different type of word stress. Word stress in Persian is purely tonic, in Tajik – mainly tonic, but in Dari it is principally quantitative, i.e. stressed syllable is longer, as a rule (analysis shows us a 10% lengthening of stressed vowel and a 35% lengthening of stressed syllable). Speech synthesis allowed us to distinguish the quantitative boundaries of that contrast (even a 10% lengthening of a syllable is sufficient to be perceived as stressed). Tone pattern appeared to be non-significant for the stress. Quantitative component of the word stress can be neutralized by a 25% tone and/or intensity rise in the unstressed syllable.

• Dari • Iranian languages • Persian • Tajik • Farsi • synthesis • prosody • stress • accent.

Введение

Иранская группа языков выделилась из индоиранской ветви индоевропейской семьи на рубеже III–II тыс. до н.э. Наиболее древним представителем этой группы является язык Авесты — религиозного памятника последователей пророка Заратуштры (середина II тыс. до н.э.). Среди других мёртвых языков иранской группы можно отметить скифский (IX–IV вв. до н.э.), мидийский (IX–VI вв. до н.э.) и древнеперсидский языки (VI–IV вв. до н.э.). Последний относится к юго-западной ветви иранских языков. Его потомками были среднеперсидский (III в. до н.э. — VII в. н.э.) и новоперсидский языки (IX–XV вв. н.э.). Современные персидский, таджикский и дари происходят от новопер-

сидского языка (рис.1). Персидский, называемый также языком фарси́, является государственным языком Ирана, таджикский — Таджикистана, дари — один из двух государственных языков Афганистана.

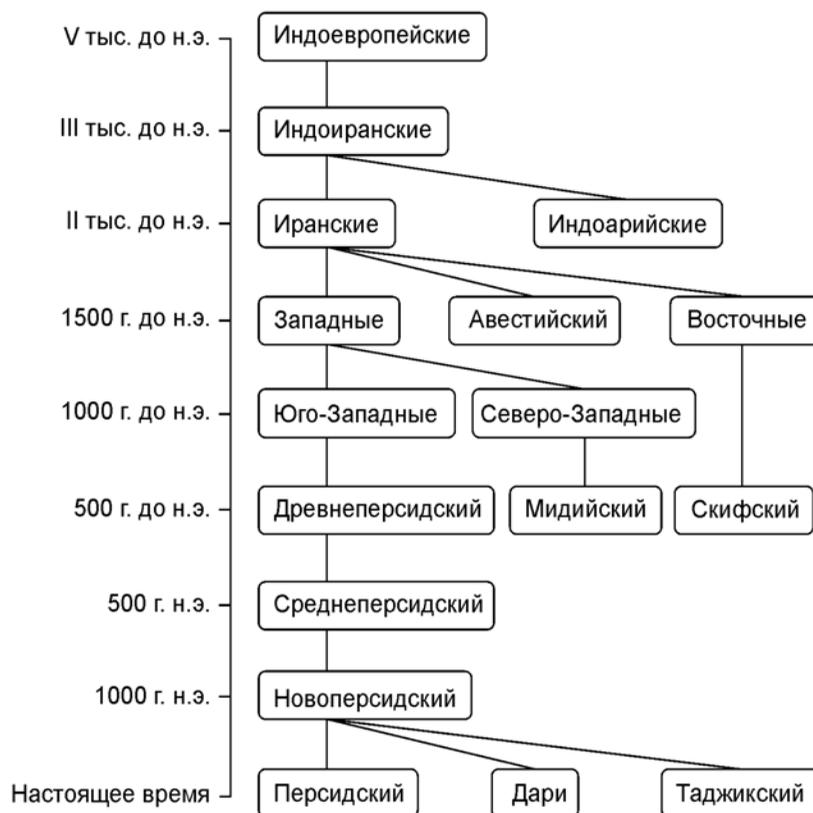


Рис. 1. Происхождение языка дари

Подавляющая часть населения Афганистана, численность которого оценивается в 32 млн. человек, как минимум двуязычна: владеет языками дари и пашто (пушту). В некоторых областях Афганистана, как, например, в Бадахшане, граничащем с таджикским Памиром, люди не говорят на пашто, но помимо дари говорят ещё на одном-двух языках (например, на шугнанском и ваханском). На языке дари, таким образом, можно общаться по всей территории Афганистана и даже за её пределами: в Иране, Таджикистане, в Узбекистане, в Пакистане, в Индии и Китае.

Само официальное название «язык дари» по отношению к современному языку Афганистана появилось исторически недавно, в афганской конституции 1965 года. Ранее этот термин относили к новоперсидскому языку. Поэтому до сих пор в повседневном общении носители языка называют его фарси́, т.е. персидский. Носители трёх близкородственных языков (персидского, таджикского и дари) прекрасно общаются между собой без переводчика. Иранская Академия языка и литературы по этой причине отказывает таджикскому и дари в статусе отдельных языков, считая их диалектами персидского. В российском и мировом востоковедении их считают различными языками, поскольку они требуют отдельного преподавания и обозначены как государственные языки в конституциях соответствующих стран.

Отличия между персидским, таджикским и дари отмечаются на всех уровнях: в фонетике, в морфологии, в лексике, во фразеологии.

Для дари на фонетическом уровне характерно:

- а) восьмифонемный вокализм с долгими [ē], [ō] против шестифонемного в персидском и таджикском языках;
- б) квазидифтонги [au], [aw] против [eu], [ow] в персидском и [au], [av] в таджикском;
- в) различие согласных фонем [q], [ɣ], которые также сохранились в таджикском, но объединились в одну фонему [q] в персидском;
- г) наличие краткого [a] в исходе слова, которое сохранилось в таджикском, но перешло в [e] в персидском;
- д) опущение согласного [h], в то время как в таджикском и персидском он сохраняется.

Есть и некоторые другие менее регулярные, но на слух хорошо воспринимаемые отличия языка дари от близкородственных. Вариативности вокализма немало способствовала арабская письменность, которая не фиксирует на письме краткие гласные. Для таджикского и дари характерно «аканье», т.е. вероятность появления в речи краткого [a] очень высока. Для персидского языка характерно «эканье».

Уже по одной формуле приветствия можно весьма точно указать место, из которого прибыл носитель языка. Для определения места проживания можно также попросить его посчитать до восьми (разговорные формы числительных приведены в таблице 1).

Таблица 1

Ряды числительных в персидском, таджикском и дари

	Дари	Персидский	Таджикский
1	yak	yek/ye	yak
2	du	do	du
3	se	se	se
4	čār	čār	čor
5	pañj	pañj	pañj
6	šaš	šeš	šaš
7	aft	haft	haft
8	ašt	hašt	hašt

Структура предложения во всех трёх близкородственных языках относится к категории SOV (субъект — объект — сказуемое). Языки хорошо различимы по интонационному рисунку. На слух европейца персидский язык наиболее певуч. В персидском повествовательном предложении на группе подлежащего происходит самое значительное повышение тона (особенно, в исфаганском диалекте), далее следует ощутимая пауза и на группе сказуемого тон идёт вниз. Глагол, находящийся в самом конце предложения, как правило, самостоятельного синтагматического (иногда словесного) ударения не принимает, т.е. произносится как клитика. В языке дари описанная интонация существует, но она не так ярко выражена. Таджикская повествовательная интонация представляется довольно ровной. Описанная повествовательная интонация в общем случае находится за пределами системы интонационных конструкций Е.А. Брызгуновой [5], так как различны типичные синтаксические структуры славянских и иранских языков (SVO vs. SOV). Данная интонация, особенно в её иранском варианте, близка к тюркской (азербайджанской, турецкой), где структура предложения также SOV. Возможно, она будет противопоставлена и более обширным инвентарям просодических признаков, выделяемых на основе русского языка [14, с. 14], что может быть использовано для автоматического различения языков типа SVO vs. SOV и требует отдельного исследования.

В персидском языке есть уникальная интонаема вопросительного предложения без вопросительного слова (переспрос), в которой тон, плавно повышаясь к концу предложения, на последнем слоге падает и вновь повышается до максимума и удерживается на этом уровне. Тонировка последнего слога переспроса сильно напоминает 3-й тон в китайском языке. Эта интонаема также отсутствует в известных нам системах интонационных конструкций русского языка. В таджикском и дари этой интонаемы нет, там тон переспроса более ровный, весьма схожий с русской интонацией.

В среде ираноязычных народностей из поколения в поколение передаётся техника чтения классических стихов, основанная на квантитативной системе противопоставления слогов (арузе) и уходящая своими корнями в первое тысячелетие н.э. Если попросить носителей прочесть отрывки из Омара Хайяма или Хафиза, то по интонационному рисунку будет практически невозможно определить их языковую принадлежность. Межъязыковые различия в стихотворной речи нивелируются, и для определения языка на первый план выдвигаются упомянутые ранее сегментные особенности. Для распознавания языка и диалекта говорящего более показательна разговорная речь, а литературная и архаичная фразеология, наоборот, затушёвывает его. Поэтому стихотворную речь можно попробовать использовать для различения групп языков, например, славянские vs. иранские.

Данная статья является продолжением исследования, результаты которого можно суммировать следующим образом: три информанта (носители языка дари) произносили различные по лексическому значению, но одинаковые по фонемному составу слова, отличавшиеся лишь местом ударения. Анализ звукозаписей показал, что соотношение длительностей слогов — наиболее информативный критерий места ударения. Тон также маркировал ударный слог, но мог быть нейтрализован вопросительной интонацией. Интенсивность не играла статистически существенной роли. В результате ударение в дари было определено как квантитативное [12].

Сопоставляя результаты этого исследования с данными других языков, отметим, что в близкородственных иранских языках — в персидском и таджикском — ударение оказалось тоническим. Причём в персидском языке оно чисто тоническое — частота основного тона (F_0) ударного слога всегда выше, чем безударного [8], а в таджикском — оно по преимуществу тоническое, т.е. F_0 в ударном слоге, как правило, выше, чем в безударном [18]. В языке пашто (другом государственном языке Афганистана, принадлежащем юго-восточной ветви иранской группы) ударение оказалось квантитативным [16]. Как известно, в русском языке ударение квантитативное [7], а в английском — тоническое [4, р. 95].

Опыт показывает, что аналитические и теоретические исследования не могут охватить всех возможных сторон явления, и даже элементарные попытки его моделирования существенно уточняют и пересматривают наше знание о нём. Обычно явление полагается хорошо изученным, если его удаётся воссоздать и управлять его развитием. Для моделирования просодической структуры слова в языке дари была проведена серия экспериментов, в которых предпринималась попытка синтезировать различные акцентные структуры для проверки гипотезы о квантитативном ударении в дари и выявления пределов вариации основных просодических параметров.

В первом эксперименте рассматривалось противопоставление двух синтезированных контрастирующих акцентных структур, отмечаемых в ряде

иранских языков. По наиболее типичным реализациям слов *راحت* *rāhat* *твой путь* vs. *راحت* *rāhāt* *спокойный* были выбраны параметры синтеза: временные значения, формантные частоты, амплитудные характеристики и тон. По ним был синтезирован исходный стимул (рис. 2).

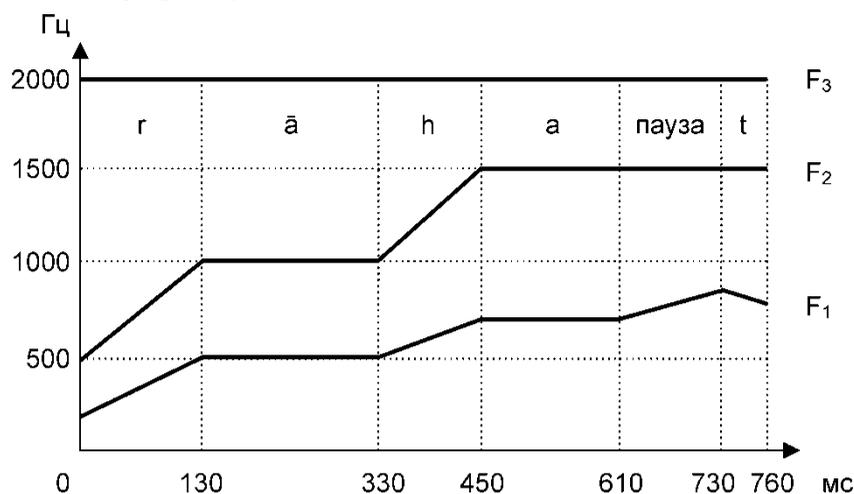


Рис. 2. Формантно-временная структура синтезируемого слова *راحت* *rāhāt* *спокойный*.

Жирными линиями обозначено движение формант F_1 , F_2 , F_3 .

Вертикальными пунктирными линиями обозначены границы сегментов

Синтез проводился с помощью программного синтезатора Д. Клатта [2] на компьютере PDP-9 в Массачусетском технологическом институте (Бостон, США) под руководством профессора К.Н. Стивенса. Синтезатор был настроен на моделирование английской речи, поэтому мы и без дополнительной настройки получили удовлетворительные звуки дари, сходные с соответствующими английскими¹. Естественно звучал глайд [h], который близок к соответствующему английскому звуку. Конечный [t] был получен из английского отключением придыхания. Афганцы сочли в дальнейшем звучание почти всех компонентов стимулов весьма естественным. Синтезатор Д. Клатта удобен для экспериментатора возможностью автоматического вычисления амплитуд формант по их частоте и ширине. Причём нам не было необходимости определять ширину формант: вполне приемлемые значения по умолчанию были заданы Д. Клаттом при написании программы. Практически для большинства звуков речи можно ограничиться описанием формантных частот.

Определённые проблемы возникали при моделировании раскатистого [r], которое отсутствует в английском, и в лаборатории не было достаточного опыта для его синтеза. Аудиторы, в целом, правильно распознавали его, но отмечали некоторый церебральный оттенок. Чтобы дать возможность аудитору привыкнуть к звучанию такого [r] и отвлечь его внимание от машинного «акцента», стимулы неоднократно воспроизводились перед началом аудирования. Перцептивная теория говорит о том, что если восприятие идёт в речевом режиме, то, несмотря на помехи, каковыми можно считать машинный «акцент», стимулы доступны лингвистической обработке [6, глава 1]. Во всяком случае, неточности конструирования [r] на сегментном уровне не влияли на общую картину просодики и интонации всего слова в целом.

Основой для конструирования континуума были наблюдения над естественной речью. Усреднённая по трём измерениям темпоральная структура реализаций, на основе которых строились серии стимулов, приведена на рис. 3. Длительность слова *راحت* *rāhat*

¹ В этом случае был проведён так называемый синтез с нуля. То есть в синтезирующую программу вводился ряд параметров в цифровом и графическом виде вручную, на основании прошлого опыта.

составила 707 мс, راحت *rāhāt* — 618 мс. Длительность каждого звука речи представлена в процентах ко всей длительности слова. Перенос ударения на первый слог удлиняет как его слогоноситель (25% → 34%), так и весь слог в целом (18% + 25% → 25% + 34%).

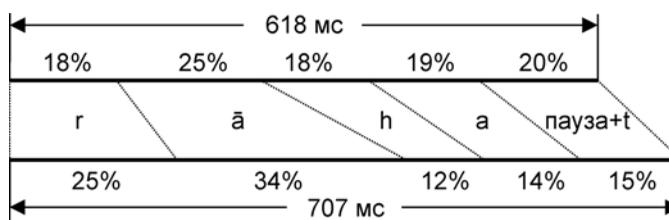


Рис. 3. Временные пропорции реализаций слов *rāhat* (внизу) vs. *rāhāt* (вверху). Пунктирные линии обозначают границы сегментов

Данный эксперимент планировался тогда, когда по более ранним исследованиям уже было известно, что в родственных языках — персидском и таджикском — ударение тоническое или по преимуществу таковое [8; 18], а ударение в дари ещё не было обследовано экспериментально. Поэтому, прежде всего, проверялась гипотеза о релевантности тона. Для этого в исходном стимуле при постоянных темпоральных и формантных структурах ступенчато варьировалась частота основного тона. Таким путём было получено 12 производных стимулов с различным соотношением тонов, тональная структура первых пяти показана на рис. 4.

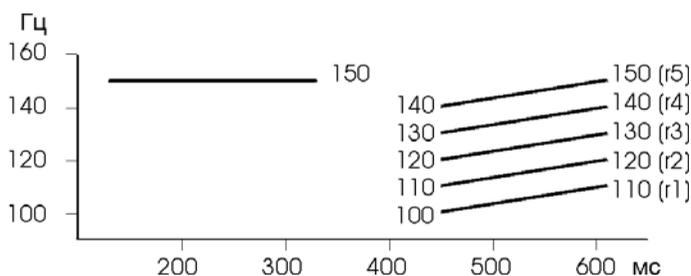


Рис. 4. Тональная структура синтезированных стимулов *r1* — *r5* на основе слова *rāhāt*

В стимулах *r1* — *r5* первые слоги одинаковы по всем параметрам. F_0 на всём их протяжении равно 150 Гц. Тон вторых слогов в каждом стимуле имеет свою собственную начальную частоту, но везде возрастает на 10 Гц.

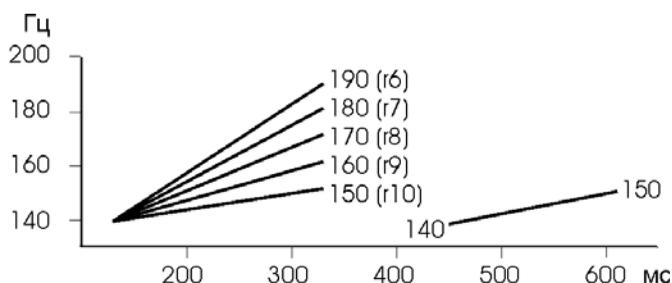


Рис. 5. Тональная структура стимулов *r6* — *r10*

Стимулы r6 — r10 различаются тоном первого слога, который повышается от 140 Гц до указанной на рис. 5 величины. Вторые слоги у них одинаковы по всем параметрам и тождественны второму слогу стимула r5.

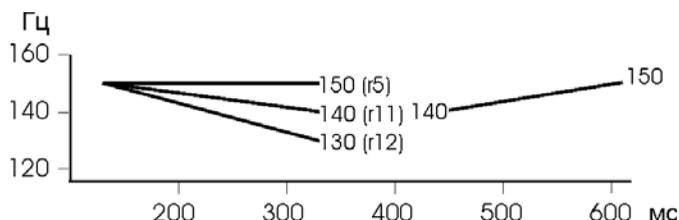


Рис. 6. Тональная структура стимулов r5, r11, r12

Стимулы r5, r11, r12 имеют идентичные вторые слоги (рис. 6) и различаются лишь тоном первого слога.

Стимулы r1 — r12 были скомпонованы в четыре аудиотеста, в каждом из которых они повторялись случайным образом пять раз. После каждого стимула была вставлена трёхсекундная пауза. После каждого десятого стимула добавлялся ещё один трёхсекундный интервал. Итого, после каждой группы из десяти стимулов пауза длилась шесть секунд. С точки зрения интонации аудиотесты интерпретировались как слова, начитанные перечислительной интонацией.

Тест 1 проводился на стимулах r1 — r5. В нём участвовало четыре информанта (все аудиторы после окончания лица в Кабуле были студентами подготовительного факультета МГУ; аудитор А3 — девушка). Перед прослушиванием идентификационного теста аудиторам ставилась задача отметить те стимулы, которые они воспринимают как слово راجت *rāhāt* *спокойный* (т.е. с ударением на последнем слоге). Число утвердительных ответов по каждому стимулу приведено в таблице 2. Дисперсионный анализ тестирования дал неоднозначные результаты.

Таблица 2

Восприятие финального ударения в стимулах r1 — r5

Стимулы/ Аудиторы	r1	r2	r3	r4	r5
A1 (С.З.)	0	0	0	0	2
A2 (Ф.К.)	3	2	1	0	5
A3 (С.Б.)	2	2	3	3	5
A4 (А.М.)	4	4	5	3	0
Итого	9	8	9	6	12

Первые три информанта А1 — А3 проявили высокую чувствительность к изменениям в просодике в зависимости от высоты тона ($p < 0,02$). Различия между аудиторами были высокозначимыми ($p < 0,0001$). Все информанты улавливали перемещение ударения на последний слог по мере повышения в нём тона (F_0), но пороговые значения F_0 , при которых ощущалось это перемещение, для каждого из аудиторов было различным.

Реже других финальное ударение «слышал» информант А1: стимулы r1 — r4, по его мнению, всегда обладали начальным ударением и только r5 в двух случаях создал эффект финального ударения. Чаще других финальную акцентуацию отмечала А3. Стимул r5 всегда относился ей к этой категории, а остальные стимулы — в 40% — 60% случаев. Мы не можем связать различия в разбросе ответов аудиторов с ка-

кими-либо образовательными или диалектными категориями, поэтому с точки зрения эксперимента считаем их случайными или относим к индивидуальным особенностям. Для этой группы аудиторов стимул r5 оказался существенно противопоставленным остальным. В нём оба слога были одинаково тонированы. Понижение тона (F_0) во втором слоге на 6,7%, как это имело место в стимуле r4, приводило к сдвигу ударения на первый слог. Дальнейшее понижение тона (как в стимулах r1 — r3) не влияло на опознавание просодики. В этой части исследования всё указывало на то, что ударение в языке дари связано с высотой тона слога (F_0).

Против ожидания, привлечение к эксперименту информанта А4 не подтвердило этот вывод.

С учётом показателей информанта А4 влияние тона (F_0) на просодику потеряло статистическую значимость². Причём помимо количественных различий обнаружились различия качественные. Вопреки мнению аудиторов А1 — А3 (рис. 7) информант А4 ощущал перемещение ударения на последний слог по мере падения тона в нём, о чём свидетельствует значимое взаимодействие факторов ($p < 0,01$; первый фактор — изменения в тоне, второй — различия в аудиторах).

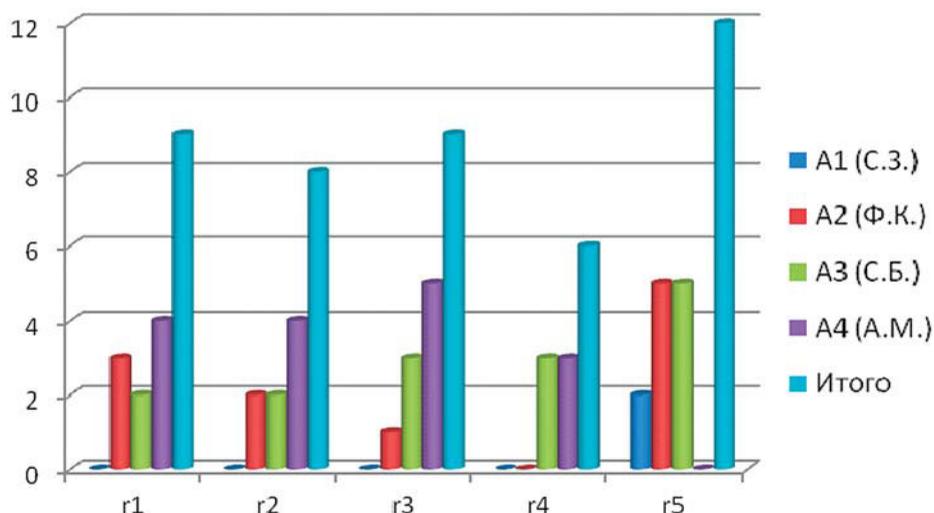


Рис. 7. Восприятие финального ударения в стимулах r1 — r5 аудиторами А1 — А4

Схожие тенденции были обнаружены в **Тесте 2**, где изучалось восприятие стимулов r6 — r10. В них ступенчато менялся тон (F_0) первого слога при постоянстве во втором. Только один аудитор А1 закономерно связывал место ударения с тоном (см. таблица 3 и рис. 8). Разброс суждений остальных, однако, полностью затушевывали эту закономерность, сделав её меньше порога значимости ($p > 0,15$). Восприятие стимулов r9 и r10 аналогично стимулам r4 и r5 в **тесте 1** оказалось статистически противопоставленным восприятию остальных стимулов, где тоны имели больший контраст.

² N.B. — редкий случай! Обычно дополнительных информантов привлекают, чтобы усилить статистическую значимость выводов, если «немного не хватает» до критического уровня (как правило, $p = 0,05$).

Таблица 3

Восприятие финального ударения в стимулах r6 – r10

Стимулы/ Аудиторы	r6	r7	r8	r9	r10
A1 (С.З.)	0	0	0	2	5
A2 (Ф.К.)	1	3	2	2	4
A3 (С.Б.)	1	1	3	5	4
A4 (А.М.)	2	2	1	1	1
Итого	4	6	6	10	14

В Тест 3 были сведены стимулы r5, r10, r11, r12, которые различаются не максимальной высотой тона (F_0), а только его контуром. Здесь проверялась гипотеза о влиянии контура на ударение. Были использованы общие, часто встречающиеся в лингвистических описаниях типы контуров: ровный, восходящий, нисходящий. Поскольку тестирование не обнаружило статистической связи этих контуров с ударением, другие, более сложные типы контуров, не рассматривались.

Тест 4 включал стимулы r14 и r15. Тон второго слога в них в отличие от всех других стимулов был выше первого: в r14 — на 10 Гц, в r15 — на 20 Гц. Половина аудиторов услышала в них финальное ударение, другая половина — начальное. Причём внутри этих подгрупп оценки информантов были достаточно хорошо согласованы, что говорит о неслучайности суждений. Таким образом, и этот тест не подтвердил гипотезу о релевантности тона (F_0) для ударения дари.

На основании результатов **Эксперимента 1** можно сформулировать следующую гипотезу. В языке дари соотношение частот основного тона начального и финального слогов так же, как и контур тона в них, с лингвистической точки зрения несущественны для распознавания просодики ударения. С психологической точки зрения некоторые носители ощущают связь этих факторов с просодикой, но эта связь очень субъективна. Одни слышат ударение на финальном слоге, если тон в нём выше. Другие воспринимают ударение на том же слоге, если тон в нём ниже. Кроме того, носители языка существенно различаются по пороговым величинам, при которых ими ощущается смещение ударения с одного слога на другой.

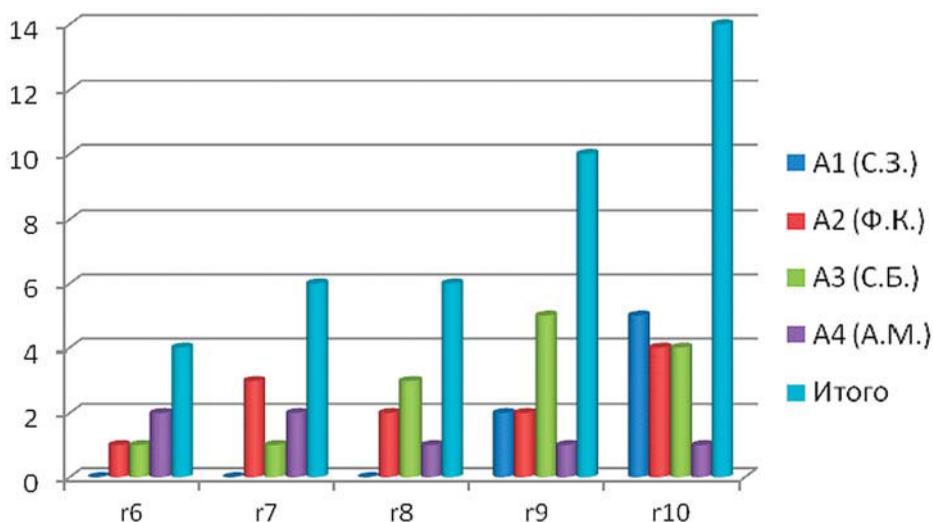


Рис. 8. Восприятие финального ударения в стимулах r6 – r10

Временная структура стимулов была получена по усреднённым данным. Значит, гипотетически возможны случаи, когда темпоральная структура вступает в противоречие с тональной. В первом эксперименте такие случаи специально не рассматривались, поэтому пока трудно судить, какой фактор имеет преимущество. Оценивая эксперимент, нужно иметь в виду, что во время тестирования аудиторы могли допускать грубые ошибки?³. Всё это предполагает дальнейшую проверку гипотез на ином материале и с другими информантами.

В **Эксперименте 2** изучался один из наиболее вероятных случаев нефинального ударения в дари — присоединение к именной части речи энклитического безударного артикля — *ē*: например, *Sēb jābloko* — *sēb-ē kakoe-to jābloko*. Артикль, аналогично другим местоименным энклитикам, входит в состав того же фонетического слова, что и имя, и не образует стыка со следующей морфемой, который мог бы сигнализировать носителю языка о словесной границе. Возможно, это обстоятельство и позволило В.И. Миколайчику считать артикль дистантным суффиксом [15, Введение]. На наш взгляд, однако, верна традиционная точка зрения на артикль как служебную часть речи, так как она согласуется с широко принятым постулатом о непроницаемости слова в иранских языках⁴. Пересмотр этого положения, конечно, возможен с формальной точки зрения, но приведёт к очень громоздкой грамматике. В частности, в случае отказа от принципа непроницаемости пришлось бы вводить целый ряд правил, объясняющих, почему словосочетания и предложения нельзя считать словами.

Исследование ударения на минимальных парах предполагает, что структуре с артиклем должна быть противопоставлена другая идентичная по фонемному составу структура, но с ударным [*ē*] в абсолютном исходе. Проблема заключается в том, что ударное [*ē*] в абсолютном исходе встречается в языке дари крайне редко. Это является ещё одним подтверждением слабой фонемной противопоставленности гласных⁵. Слабость, которая компенсируется закреплённостью гласных за определёнными слоговыми структурами и позициями по отношению к ударению. Односложные слова типа *bē bez*, *pē net*, *čē čo*, *sē tri* в счёт не идут, так как в них нет безударных слогов и поэтому их акцентуацию можно считать финальной с таким же успехом, как и начальной.

Единственной подходящей для просодического экспериментирования структурой можно считать слово *šanbē subbota* и производные от него *yakšanbē voskresen'je*, *došanbē ponedel'nik* и т.д. до *ranjšanbē četverg*. К сожалению, для корректности эксперимента в языке отсутствует слово *šanb**, которое в сочетании с артиклем создавало бы контрастную структуру с начальной акцентуацией. Поэтому вновь пришлось обратиться к искусственным словам, и за исходный материал было взято нонсенс-слово

³ Это замечание, вероятно, можно отнести к информанту А4 (его ответы статистически значимо отличались от ответов остальных). Под грубой ошибкой мы понимаем опечатку, когда информант просто не там поставил галочку. В дальнейшем могло произойти закрепление неправильной реакции, что могло дать значительную систематическую ошибку.

⁴ Полагается, что между морфемами одного слова нельзя вставить другое слово, в частности, между корнем и суффиксом. Если же считать артикль суффиксом, то придётся мириться с тем, что между корнем существительного и его суффиксом можно вставить прилагательное. Более подробно об этой дискуссии [см. 9].

⁵ Как было показано в работе [10], все 8 гласных языка дари разбиваются на 3 кластера: [*i-ē-e*], [*a-ā*] и [*o-ō-u*], внутри которых в значительном слое лексики они могут в зависимости от стиля речи переходить друг в друга, не меняя значения слова. В этом и заключается их слабая фонемная противопоставленность.

bēbē*. В этом слове два одинаковых по фонемному составу слога, параметры которых можно варьировать при синтезе независимо друг от друга. Структура с ударением на последнем слоге соответствует именно типу акцентуации, которое должно восприниматься носителями как некое существительное или прилагательное. Структура с нефинальным ударением выглядит как существительное bēb* с артиклем -ē, т.е. bēb-ē*.

Синтез проводился на программном синтезаторе Д. Клатта, реализованном экспериментатором на ЭВМ СМ-4 Филфака МГУ. Для воспроизведения звука использовался комплекс КАМАК с цифро-аналоговым преобразователем, который управлялся специальной программой [17]. Адаптация синтезатора к комплексу СМ-4 КАМАК потребовала существенных изменений в программе Клатта. Кроме того, нам потребовалось написать программу генерации идентификационных тестов, совместимую с использованной ранее американской в **Эксперименте 1**.

Гласные слова bēbē были идентичны по формантному составу: три низшие форманты имели следующие частоты: $F_1 = 450$ Гц, $F_2 = 2200$ Гц, $F_3 = 2800$ Гц. Их ширина соответственно равнялась 60, 150 и 300 Гц. Эти величины были подобраны, исходя из результатов эксперимента по восприятию гласных иранцами и таджиками [11], таким образом, чтобы синтезированный [ē] звучал как нечто среднее между [e] и [i] персидскими⁶.

Длительность первого слога, параметры которого во всех стимулах были одинаковы, составляла 200 мс. Согласный [b] был создан формантными переходами в начале слога без привнесения дополнительных шумовых характеристик. Он занимал 30 мс. За это время частоты формант F_2 и F_3 линейно возрастали на 500 Гц до указанных выше величин, т.е. до квазистационарного участка гласного. Частота первой форманты (F_1) за это время возрастала от 150 до 450 Гц, т.е. на 300 Гц.

Амплитуда голосового возбуждения в течение первых 30 мс возрастала от 0 до 60 Дб, затем поддерживалась постоянной в течение 140 мс, а за последние 30 мс звучания слога уменьшалась до 0. Затем следовала 50-миллисекундная пауза, которая разделяла слоги (см. рис. 9).

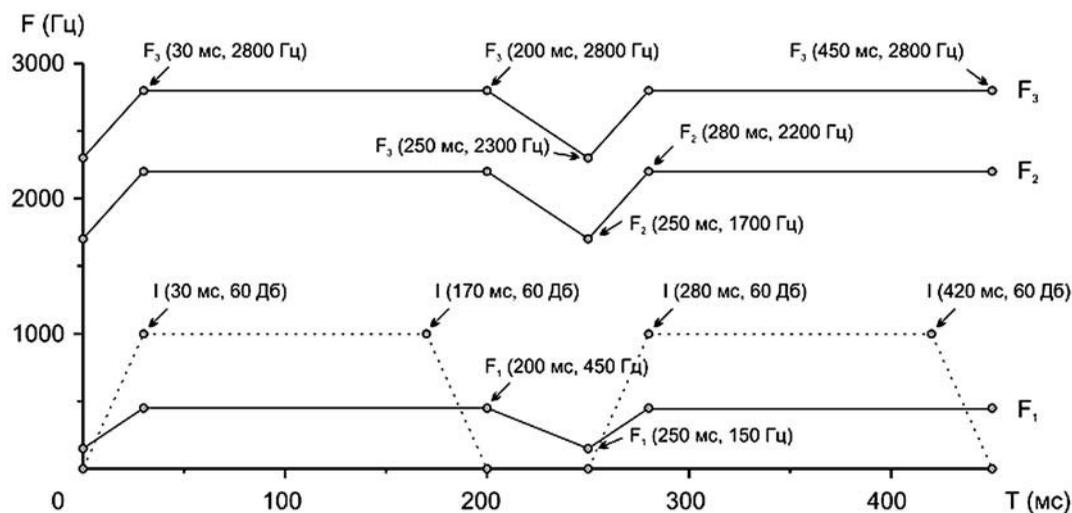


Рис. 9. Формантно-временная структура базового стимула nonsens-слова bēbē. Контур интенсивности (I) обозначен пунктиром

⁶ Это помещает его в середину кластера [i-ē-e], где он должен теоретически находиться с точки зрения артикуляции.

Работа синтезатора предполагает использование и других параметров помимо описанных, которые были взяты из статьи [2, р. 976]. Параметр GO (общее усиление) в каждом стимуле подбирался экспериментально для устранения искажений при оптимальном соотношении «сигнал/шум». Различия в громкости стимулов не превышали 5 Дб. Синтезатор Д. Клатта моделирует инерционность речевого тракта, поэтому при изменении амплитуды голосового возбуждения интенсивность сигнала меняется не мгновенно, а за время порядка 5–10 мс. Другая особенность — моделирование взаимосвязи интенсивности с частотой основного тона, т.е. при прочих равных условиях стимулы с более высоким тоном получаются более громкими. Эти особенности обуславливают частные отклонения интенсивности стимулов от контура на рис. 9. По оси X откладывается время в мс, по оси Y — частота в Гц. Пунктирный контур (I) демонстрирует амплитудные характеристики, т.е. наличие фонации.

В эксперименте проверялись гипотезы о влиянии интенсивности (I), длительности (T), частоты основного тона (F_0) и его контура на восприятие ударения. Первые три фактора варьировались только во втором слоге, контур менялся в обоих слогах. Итого, варьировалось пять факторов, каждый из которых мог принимать два значения. Количественные факторы — интенсивность (I), длительность (T) и тон (F_0) — могли быть во втором слоге либо на 25% выше, чем в первом, либо на столько же процентов ниже. Уровень в 25% был взят как заведомо более высокий, чем тот, который бывает в естественной речи, чтобы, по выражению специалистов по планированию экспериментов, обязательно накрыть критические точки зависимостей. Качественный фактор — контур тона — мог быть либо падающим, либо восходящим. В том и другом случае приращение составляло 20 Гц на всю длительность слога, что тоже заведомо выше порога чувствительности информантов, который для людей с нормальным (не тренированным) слухом составляет около 5 Гц.

Варьирование пяти факторов на двух уровнях даёт нам 2^5 комбинаций, т.е. 32 стимула. Эти стимулы были сведены в идентификационный тест, в котором они случайным образом повторялись два раза со стандартными паузами (см. описание **Эксперимента 1**). В отличие от экспериментирования со словом راحت *rāhāt* *удобный*, когда перед аудиторами ставилась привычная задача соотнести звучание с написанием, анкета второго эксперимента потребовала более сложной установки. Информантам A5 и A6 ставилась задача распознать слово с артиклем (و حبت ی *uā-ye wahdat*), что предполагает их несколько более высокую лингвистическую образованность, не превышающей, однако, требований лица. В анкете был напечатан ряд нулей и единиц. Аудиторам было объяснено по-персидски, что при опознавании слова с артиклем они должны отметить единицу, в противном случае — ноль.

Для проверки усвоения объяснений перед тестированием было проведено две тренировки. Крайние точки континуума, т.е. стимул с явным ударением на начальном и стимул с явным ударением на финальном слоге, были сведены в отдельный аудиотест, где они повторялись в случайном порядке десять раз. После каждой тренировки анкеты анализировались, обсуждались с аудиторами и в случае ошибок делались соответствующие замечания. Информанты хорошо усваивали поставленную задачу и уже после первой тренировки грубых ошибок не допускали.

Дисперсионный анализ показал, что все три количественных фактора — длительность (T), интенсивность (I) и тон (F_0) — существенно связаны с просодикой ударения (все ниже критического уровня $p = 0,05$). Наиболее зна-

чимым оказался тон ($p = 0,005$), затем следуют интенсивность ($p = 0,01$) и длительность ($p = 0,03$). Применённая методика дисперсионного анализа не позволяла, однако, определить степень значимости различия между факторами. Контур тона в слогах оказался незначимым, что подтвердило выводы предыдущего эксперимента.

Описываемый тест, как и предыдущий эксперимент, выявил определённую склонность информантов к более частому восприятию финального ударения. Во всех стимулах, где три фактора возрастали во втором слоге, слышалось финальное ударение. Но только 75% реализаций с обратным соотношением параметров воспринято с начальным ударением. Мы можем объяснить эту асимметрию большей распространённостью финального ударения в дари, так что требуется дополнительное усиление акустических факторов, чтобы побудить информантов услышать нечто противоречащее их ожиданиям⁷.

Особый интерес представляют случаи, когда факторы действуют противоречиво, т.е. два из них выделяют первый слог, один из них — второй. И здесь эффект финального ударения достигается значительно легче, чем начального. Тонем и интенсивностью по отдельности удаётся превозмочь действие остальных факторов в 58,3% случаев. Противодействие длительности выражено значительно слабее: только в 1/3 предъявлений была воспринята финальная акцентуация.

В противоположной части континуума, там, где первый слог выделялся одним фактором, второй — двумя, наибольшее противодействие восприятию финального ударения оказывала длительность (16,7%). Более высокий тон первого слога только в 8,8% нейтрализовал остальные факторы. Большая интенсивность первого слога ни разу не помешала аудиторам воспринять последний слог ударным.

В целом, результаты **Эксперимента 2** можно интерпретировать следующим образом. При значительном возрастании (около 25%) длительности, интенсивности и тона в одном из слогов каждый фактор становится значимым и связывается информантами с просодикой слова. Для однозначной интерпретации типа акцентуации достаточно согласованного воздействия двух факторов из трёх. Оставшийся фактор может противодействовать первым двум, практически не влияя на ударение.

При прочих равных условиях чаще слышится ударение на финальном слоге. Это хорошо согласуется с нашими предыдущими исследованиями и с общеизвестной гипотезой о более сильном выделении начального слога. В самом деле, обычно в естественной речи первый слог характеризуется более высокими уровнями факторов при изолированном произнесении слова. Когда же оба слога одинаковы и аудиторы не слышат усиления инициала, они полагают, что выделена финаль. Используемый уровень различия слогов $\pm 25\%$, очевидно, не предельно, что предполагает возможность синтеза слов с ещё большими межслоговыми различиями.

Эксперимент 3. Ранее обследованные модели просодических оппозиций типа «имя + энклитика», где в качестве энклитики может быть местоимение или артикль, в противопоставлении цельному слову применимы ко многим иранским языкам. В трёх близкородственных языках — персидском, таджикском и дари — эти модели охватывают значительную часть случаев нефинального ударения в любом тексте.

Существует, однако, специфический для дари случай нефинальной акцентуации. Он обнаруживается тогда, когда заимствованное из арабского языка слово оканчивается на «согласный + 'айн». Такие слова даже в полном стиле произносятся так, как будто вместо 'айна звучит гласный [a]. Ударение не падает на него, так как его основной фонологический вариант является согласным. Например, *góba* из رُبْع *rob'* *четверть*, *tába* из طَبْع *tab'* *печать*. При образовании словосочетаний, когда после такого рода слов появляется морфема с гласным началом, гласный [a] исчезает или заменяется

⁷ Это явление аналогично перцептивной защите, которая проявляется, когда частота опознавания неприличных слов становится существенно меньше, чем приличных.

согласным [ʔ] — перерывом в голосе. Например, *طبع tába печать*, но *اول طبّع tabʔ-e awwá I первое издание*.

К сожалению, и в этом случае не удаётся обнаружить чистых словарных оппозиций. Попытка использовать оппозицию *قلع qála олово* — *قلع qalá крепость* не удалась, так как произношение *qalá* характерно только для неполного стиля. В полном же стиле, в котором информанты привыкли начитывать тексты, это слово произносится *qaʔa*. Поскольку из-за смешения стилей от информантов не удалось добиться естественного произношения, эта оппозиция не рассматривалась.

В отличие от **Эксперимента 2**, где слоги резко противопоставлялись друг другу (параметры изменялись скачком в $\pm 25\%$), в **Эксперименте 3** были запланированы более тонкие различия. В тесте 1 длительность (Т), интенсивность (I) и тон (F_0) варьировались во втором слоге в пределах $\pm 10\%$. Ранее было взято всего два уровня для каждого фактора, здесь же число уровней было доведено до пяти, в том числе, каждый параметр второго слога мог принимать значение, равное первому слогу (табл. 4). Варьирование трёх параметров на пяти уровнях даёт $5^3 = 125$ комбинаций.

Параметры первого слога во всех стимулах были постоянными и равнялись значениям, указанным в средней (третьей) строке таблицы. Таким образом, каждый параметр второго слога мог принимать значения от -10% до $+10\%$ от значения первого. Шаг в 5% процентов был выбран как весьма близкий к порогу восприятия.

Таблица 4

Уровни факторов во втором слоге нонсенс-слова *táta*

	Уровень	Интенсивность (Дб)	Тон (Гц)	Длительность (мс)
1	+10%	66	143	220
2	+5%	63	136	210
3	0%	60	130	200
4	-5%	57	124	190
5	-10%	54	117	180

Для афганских аудиторов (в то время студентов ТашГУ) были подготовлены анкеты, напечатанные в привычной для них арабо-персидской графике. В них в несколько колонок были напечатаны два опознаваемых слова: *تتە táta* (буква *ع ayn* на конце означала безударный гласный [a]) и *تتە tatá* (буква *ه hā-ye havváz* на конце обозначала ударный гласный [a]). Как и в **Эксперименте 2**, оба предъявляемых слова были нонсенс-словами.

Тест 1 был прослушан двумя аудиторами — А6 и А7. Поскольку тест не выявил влияния длительности на ударение, а влияние интенсивности и тона было близко к критическому уровню, планирование эксперимента было несколько изменено. Уровень варьирования длительности второго слога был повышен до $\pm 18\%$ с шагом в 9%, с тем чтобы подобрать такие значения, которые могли бы нейтрализовать тон. Так был получен **Тест 2**.

Тест 2 был прослушан аудиторами А7 и А8. Аудитор А8 показал очень высокую чувствительность к тону ($p < 0,001$) и интенсивности ($p < 0,0001$). Однако в его восприятии не обнаружилось существенной связи длительности (Т) с просодикой. Совместная обработка данных по обоим аудиторам не выявила значимого влияния ни одного из факторов в отдельности на вос-

приятие ударения. Обнаружились существенные расхождения в оценках аудиторов ($p < 0,001$). Тем не менее, отмечены статистически значимые взаимодействия факторов: длительности (Т) и тона (F_0 , $p < 0,02$), интенсивности (I) и тона (F_0 , $p < 0,05$). Это можно интерпретировать как то, что для создания эффекта ударения на нужном слоге недостаточно каждого фактора в отдельности, необходима их оптимальная комбинация. В тех случаях, когда у обоих слогов все параметры были идентичны, преимущества какого-либо слога выявлено не было: в половине случаев ударным указывался первый слог, в остальных — второй.

В итоге проведения **тестов 1 и 2** не удалось подобрать такого сочетания параметров, при котором длительность слогов стала бы значимой для восприятия ударения. Поэтому был запланирован **Тест 3**. В его стимулах соотношение слогов по длительности (Т) варьировалось более крупными шагами по 15 мс, т.е. по 10% от длительности каждого слога. Тон (F_0) и интенсивность (I) изменялись в более узких пределах. Приращение интенсивности (I) составляло 2 Дб, т.е. 3,3%; приращение тона (F_0) — 4 Гц, т.е. 3%. Для более естественного звучания длительность (Т) каждого слога была уменьшена до 150 мс (исходный уровень). Пауза между слогами составляла 50 мс. Так же, как и в предыдущих текстах, три параметра варьировались в различных сочетаниях на пяти уровнях, что даёт $5^3 = 125$ сочетаний.

Тест 3 был прослушан аудиторами А9 и А10. Самым существенным фактором, определяющим ударение, была длительность (Т) ($p < 0,1$). Влияние тона было далеко от критического ($p < 0,25$). Однако при незначительном увеличении числа информантов вполне возможно усиление этого фактора до значимого⁸. Влияние интенсивности (I) оказалось столь невелико, что его вероятность не поддавалась определению по статистическим таблицам.

Эксперимент 3 показал, что при небольших уровнях варьирования тона и интенсивности на первый план выходит длительность слогов как фактор, определяющий место ударения.

Заключение

Инструментальные исследования естественной речи носителей языка дари показали статистическую значимость длительности слога для восприятия места ударения. Эксперимент с синтезированной речью подтвердил этот вывод. 10-процентное увеличение длительности какого-либо слога по сравнению с другим при прочих равных условиях достаточно для восприятия его ударным.

Синтез речи позволил получить и ряд принципиально новых результатов. Выяснилось, что 20-процентное повышение тона способно нейтрализовать действие темпорального фактора. Повышение интенсивности также может нейтрализовать воздействие темпорального фактора на восприятие, но при ещё более высоких уровнях — от 25% и выше. Причём взаимосвязь тона и интенсивности, с одной стороны, и ударения, с другой, ощущается не всеми носителями дари. Изменение контура тона не влияет на восприятие ударения.

Такое соотношение просодических характеристик может быть объяснено следующими факторами:

1. Ударный слог главным образом маркируется большей длительностью. Однако более высокий тон и интенсивность могут воздействовать на восприятие ударения в каком-либо слоге как общечеловеческие психоакустические признаки затраты больших усилий на произнесение того или иного слога [3, р. 231].

⁸ Существует мнение, что при неограниченном увеличении информантов любой фактор можно представить статистически значимым. Поэтому, как это ни парадоксально, наибольшего доверия вызывают высокозначимые результаты, полученные на малых и средних выборках.



2. Принимая во внимание, что в близкородственных языках — таджикском и персидском — ударение тоническое, можно предположить, что просодическая роль тона в дари была отодвинута на второй план его историческим развитием и контактами со вторым официальным языком пашто.

За последние десятилетия на дари в меньшей степени действовали консервативные факторы в виде средств массовой информации и стандартизированного школьного образования, чем на соседние государственные — персидский и таджикский. Дари отличается повышенной вариативностью, что скорее приводит к расшатыванию произносительной системы, чем к её консолидации. Существенное влияние на дивергенцию просодических моделей персидского и таджикского, с одной стороны, и дари, с другой стороны, могли оказать контакты с языком пашто, ударение в котором имеет преимущественно квантитативный характер.

Литература

1. *Hirst Daniel and Di Cristo Albert. A survey of intonation systems // Intonation Systems: A Survey of Twenty Languages. Cambridge: Cambridge University Press, 1998. Vol. 1: 1.*
2. *Klatt D. H. Software for a cascade/parallel formant synthesizer // Journal of the Acoustical Society of America. 1980. Vol. 67(3).*
3. *Lehiste I. Suprasegmental features of speech. Contemporary issues in experimental phonetics. New York, London: Academic Press, 1976.*
4. *Roach P. English Phonetics and Phonology. A practical course. [s.l.]: Cambridge University Press, 2002.*
5. *Брызгунова Е.А. Звуки и интонация русской речи. Лингафонный курс для иностранцев. М.: [б.н.], 1969.*
6. *Галунов В.И. и Родионов В.Д. Моделирование процессов передачи информации в звуковом диапазоне. Ленинград: Внешторгиздат, 1988.*
7. *Златоустова Л.В. Фонетическая природа русского словесного ударения. Спб.: МГУ, 1953.*
8. *Иванов В.Б. Акустические характеристики персидского словесного ударения. Автореферат кандидатской диссертации. М.: МГИМО, 1976.*
9. *Иванов В.Б. Границы слова и инкапсуляция в персидском, таджикском и дари // Вопросы языкознания. М.: [б.н.], март 1995 г. С. 107–117.*
10. *Иванов В.Б. Динамика вокализма в персидском языке и дари // Исследования по иранской филологии / ред. Островский Б.Я. М.: Кафедра иранской филологии ИСАА при МГУ, 1997. Т. I: III.*
11. *Иванов В.Б. О сопоставлении персидского и таджикского вокализма на основе синтеза речи // Филологические науки. М.: МГУ, 1987(5) г.*
12. *Иванов В.Б. Ударение в дари (этап анализа) // Актуальные проблемы строя иранских языков / ред. Лебедев К.А. и Островский Б.Я. М.: МГУ, 1988.*
13. *Кодзасов С.В. и Кривнова О.Ф. Общая фонетика. М.: РГГУ, 2001. 592 с.*
14. *Кодзасов С.В. Исследования в области русской просодии. М.: Языки славянских культур, 2009. 492 с.*
15. *Миколайчик В.И. Основы теоретической грамматики персидского языка. М.: Военный университет, 1980.*

16. *Тарбеева Н.М.* Экспериментальное изучение словесного ударения в языке пашто. М.: ИСАА МГУ, 2010.
17. *Трунин-Донской В.Н. и Чучупал В.А.* Цифровая обработка речевого сигнала // Автоматическое распознавание слуховых образов (АРСО). Киев: [б.н.], 1977.
18. *Хаскашев Т.Н.* Фонетическая природа словесного ударения в современном таджикском литературном языке (экспериментально-фонетическое исследование). Спб.: ЛГУ, 1972.

Сведения об авторе

Иванов Владимир Борисович —

заведует кафедрой Иранской филологии Института стран Азии и Африки (ИСАА) МГУ им. М.В. Ломоносова, а также в качестве совместителя является ведущим научным сотрудником в секторе Иранских языков Института лингвистики РАН, доктор филологических наук, профессор. Преподаёт персидский язык, его теорию и общелингвистические дисциплины. Ведёт полевые и экспериментальные исследования в области иранских языков и диалектов. Опубликовал более 85 работ. Автор учебников по персидскому и таджикскому языкам.