

УДК 591.582.2

**АКУСТИЧЕСКИЙ И АУДИТОРСКИЙ АНАЛИЗ
ГЛАСНОПОДОБНЫХ ЗВУКОВ СЕРОГО (*PSITTACUS ERITHACUS*)
И ВОЛНИСТОГО (*MELOPSITTACUS UNDULATUS*) ПОПУГАЕВ**

© 2006 г. К. О. Уплисова

Санкт-Петербургский государственный университет 199034, Санкт-Петербург, Университетская наб., 7/9

E-mail: sehmet@fromru.com

Поступила в редакцию 11.04.2006 г.

С целью выявления стабильных акустических характеристик различных фонетических категорий гласных звуков, имитируемых говорящими птицами, были исследованы гласноподобные звуки серого и волнистого попугаев. Ранее была показана невозможность разделения категорий гласных звуков, имитируемых говорящими птицами, на основании значений частот первого и второго спектрального максимума (Patterson, Pepperberg, 1994; Уплисова, 2004), что характерно для человеческой речи. В данной работе был использован признак зависимости отношений амплитуд спектральных компонентов от их частоты, применяемый для гласноподобных звуков младенцев (Куликов и др., 1999). Результаты анализа показали, что зависимость отношений амплитуд от частоты спектральных компонентов фонетических категорий “а”, “о”, “у”, “и” для серого и волнистого попугаев различны.

Ключевые слова: имитация, говорящие птицы, попугай, гласные звуки.

ВВЕДЕНИЕ

Изучение имитационных сигналов говорящих птиц представляет особый интерес в связи с принципиальными отличиями в строении голосового аппарата и нервного обеспечения звукопродукции у птиц и человека. Тем не менее птицы различают категории речевых сигналов человека (Dooling et al., 1995) и имитируют слова, несмотря на акустические отличия голосов окружающих их людей. Имитационные звуки птиц в свою очередь распознаются человеком, хотя не могут иметь такие же акустические характеристики, что и в речи человека. На данный момент существует несколько работ, в которых были проанализированы имитации человеческой речи различными говорящими птицами.

В одной из первых работ было проведено сравнение акустических характеристик имитаций скворца-майны и голоса обучавшего птицу человека (Klatt, Stefanski, 1974). Было показано, что, хотя в сиринксе (нижней гортани) майны, как и у всех воробьиных, присутствует два независимых источника колебаний, по одному в каждом бронхе, она может достаточно точно копировать частоту основного тона и форманты хояина.

В детальном исследовании акустической структуры гласноподобных звуков двух попугаев (Patterson, Pepperberg, 1994) полученные данные соответствовали морфологическим данным и говорили о том, что попугаи обладают только од-

ним источником основного тона (Gaunt, Gaunt, 1985). Таким образом, их способ звукогенерации может считаться гораздо более близким к человеческому, чем у майны. Но, тем не менее, хотя авторы смогли определить и частоту основного тона и частоты формант гласноподобных звуков попугаев, эти значения совпали с частотами формант гласных звуков человеческой речи далеко не для всех категорий (исследовались десять категорий гласных английского языка). Из этого следовало, что основной признак, по которому традиционно разделялись различные категории гласных звуков человеческой речи, а именно значения частот первой и второй форманты, оказался неприменим для гласноподобных звуков попугаев.

Несколькими годами позже была проанализирована акустическая структура различных звуков волнистых попугайчиков (Banta Lavenex, 1999) – самых маленьких из способных к “говорению” представителей попугаеобразных. На основании детального изучения акустических характеристик (как видовых сигналов, так и имитационных звуков речи) исследовательница выдвинула предположение, что одним из механизмов звукогенерации для этих птиц является амплитудная модуляция – результат нелинейного взаимодействия двух источников колебаний (несущего и модулирующего). В результате амплитудной модуляции в спектре звука появляются частоты, не отражающие характеристики источника. Модуляция мо-

жет быть периодической или непериодической, что может приводить к гармоническоподобному спектру в первом случае и шумоподобному во втором. К звукам, образованным подобным образом, в соответствии с классическими представлениями, не применимы понятия частоты основного тона и гармоник. Механизм образования амплитудной модуляции на данный момент не известен. Вместе с тем часть вокализаций волнистого попугайчика имела и четкую гармоническую структуру.

Таким образом, на данный момент не ясно, на основании каких признаков человек может различать фонетические категории гласных, имитируемые говорящими птицами. В связи с этим была начата работа по выявлению акустических признаков стабильно присутствующих у различных категорий имитационных гласных. Были исследованы произнесения человеческой речи скворцом-майной Чикой и серыми попугаями Ромой и Кузей (Уплисова, 2004). Имитации у всех птиц были насыщены шумовыми компонентами. Звуки со строго гармонической структурой имели место у майны, и в редких случаях у двух серых попугаев. В большинстве случаев гласноподобные звуки серых попугаев имели шумоподобную структуру, поэтому частоту основного тона было невозможно выделить. Имитации слов у всех птиц были достаточно четкими, но идентификация гласного звука после выделения его из слова во многих случаях была неоднозначной.

Для всех трех птиц было показано, что на основании частотных значений двух наиболее выраженных спектральных максимумов различные категории гласноподобных звуков разделить невозможно в связи с большим разбросом значений. Поэтому был использован критерий, полученный ранее для гласноподобных вокализаций младенцев, звуки которых не разделялись на основании значений формант в связи с высокой частотой основного тона (Куликов и др., 1999). Характер распределения зависимости амплитудных значений от частоты спектральных компонентов позволил четко разделить категории "а", "о" "и" для скворца-майны и серого попугая Ромы, но для Кузи различия оказались не столь явными.

С целью определения стабильности выявленного признака различных категорий гласноподобных звуков у различных особей, относящихся к одному и разным видам, в данной работе проведено исследование гласноподобных звуков серого попугая Кузи и волнистого попугайчика Геши.

МАТЕРИАЛ И МЕТОДИКА

Запись вокализаций краснохвостого серого попугая Кузи, принадлежащего А.А. Тихомирову, была проведена на магнитофон SONY TC-D5 Pro II (неравномерность частотной характеристики ± 3 дБ в диапазоне 40–17000 Гц) с микрофоном AKG D310 в домашних условиях. Запись вокализаций волнистого попугайчика Геши, принадлежащего семье Алексеевых, осуществлялась при помощи магнитофона Marantz PMD222 с микрофоном Sennheiser E855. Оцифровка записей осуществлялась с использованием 16-разрядной звуковой карты Creative Labs AVE 64 и Aardvark DirectMix USB с частотой дискретизации 44100 Гц.

Из всего вокального репертуара птиц выделялись имитации слов человеческой речи. Помимо них в записи присутствовали видовые сигналы и имитации механических звуков. Выделение гласных звуков осуществлялось на слух, на основании равномерности звучания, по однородности динамической спектрограммы и амплитудно-динамической формы сигнала. После изоляции из состава слова гласноподобный участок соотносился исследователем с одной из категорий гласных русского языка. Если звук после выделения терял все признаки гласноподобности, в дальнейшем он не анализировался.

Спектральный анализ звуковых сигналов осуществлялся на основе быстрого преобразования Фурье, данные взвешивались при помощи окна Хемминга с 1024 фильтрами.

Так как спектры гласноподобных звуков в большинстве случаев не являлись гармоническими, то определялась частота и амплитуда всех спектральных компонентов, необходимых для описания контура огибающей. По возможности, на основании амплитудно-динамической формы, выделялась несущая и модулирующая частоты анализируемого звука.

Аудиторскому анализу были подвергнуты звуки, отнесенные исследователем как к категориям гласных "а", "о", "у", "и", так и промежуточным. Серия звуков, предназначенная для прослушивания аудитором, состояла из пятидесяти произнесений, каждое из которых повторялось по 3 раза с интервалом 400 мс, различные звуки были разделены интервалом в полторы секунды. Серии представлялись различным группам студентов, ранее не слышавшим эти звуки. При демонстрации гласноподобных звуков Геши предварительно проигрывался небольшой участок аудиозаписи "говорения" попугая. Перед аудитором ставилась задача обозначить, какой гласноподобный звук (любой) они слышат, или поставить прочерк, если предъявляемый звук нельзя было отнести ни к одной категории. Было получено двенадцать вариантов ответа (а, о, у, и, ы, е,

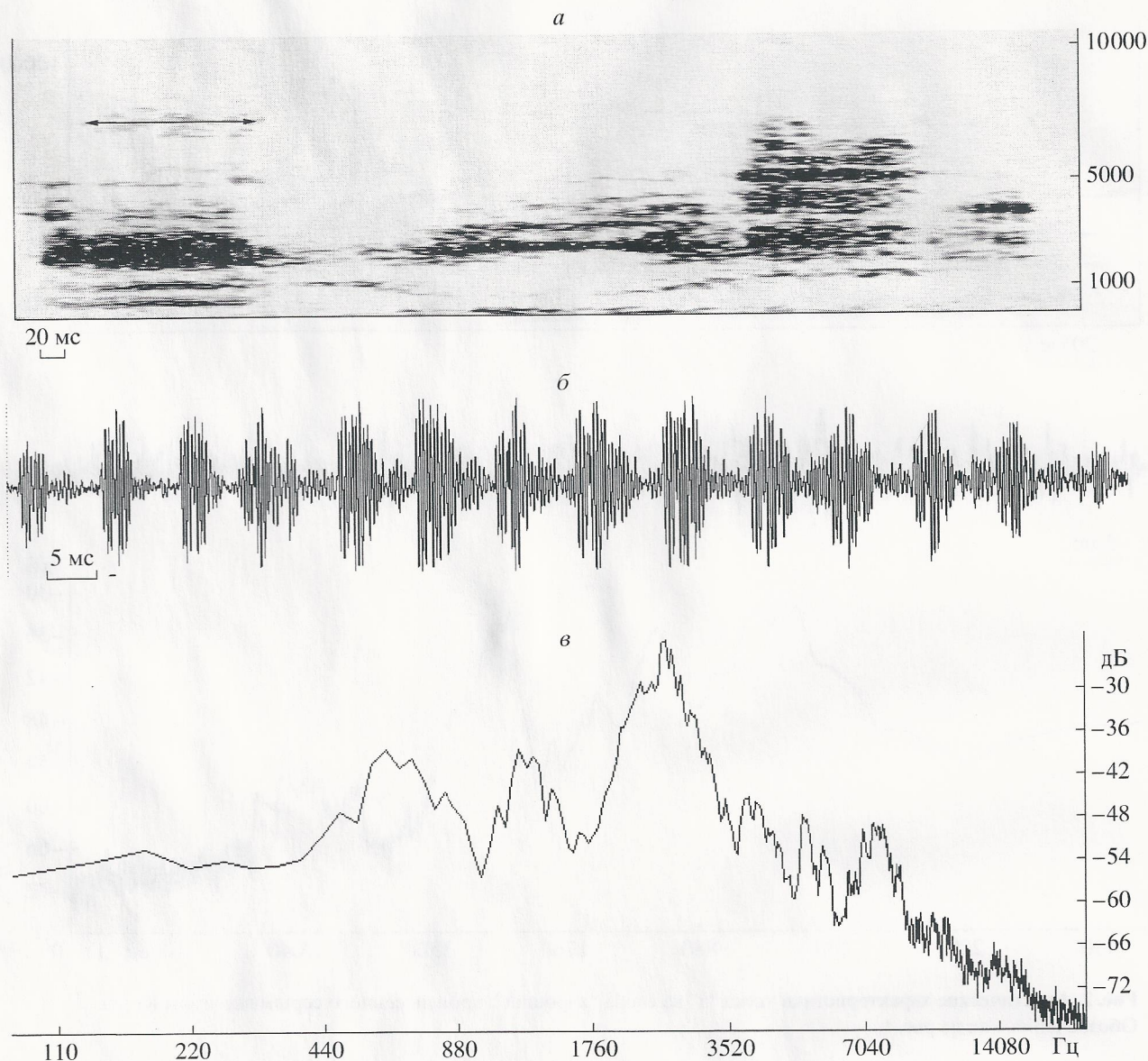


Рис. 1. Акустические характеристики звука “а” из слова “хороший”, произнесенного волнистым попугаем Гешей.
а – Динамическая спектрограмма (512 фильтров) произнесения слова “хороший”. Стрелкой выделен участок, соответствующий гласному “а”. По оси абсцисс – время, мс; по оси ординат – частота, Гц. *б* – Амплитудно-динамическая форма гласного “а”. По оси абсцисс – время, мс; по оси ординат амплитуда, отн. ед. *в* – Суммарный спектр (1024 фильтра) гласного “а”. По оси абсцисс – частота, Гц; по оси ординат – амплитуда, дБ.

ю, я, э, ё, неоднородный звук, не гласноподобный звук). Достоверность отнесения звука к одной из категорий определялась по биномиальному критерию.

На основе полученных значений амплитуд и частот спектральных составляющих проводили следующую обработку: у каждого звука последовательно из амплитуды всех зафиксированных спектральных компонентов вычиталась амплитуда первого, превысившего уровень шума. Полученная разность в децибелах переводилась в абсолютную величину. Далее значения разницы ам-

плитуды спектральных компонентов всех звуков, отнесенных к единой фонетической категории, на основании частоты распределяли по группам, соответствующим критическим полосам слуха человека (Цвикер, Фельдкеллер, 1965) и уже в каждой группе рассчитывалась медиана. Полученное значение переводили в децибелы, и вводили поправку на аудиограмму для каждой критической полосы. После этого строили распределения зависимости медиан разности амплитуд от частоты спектральных компонентов. Достоверность

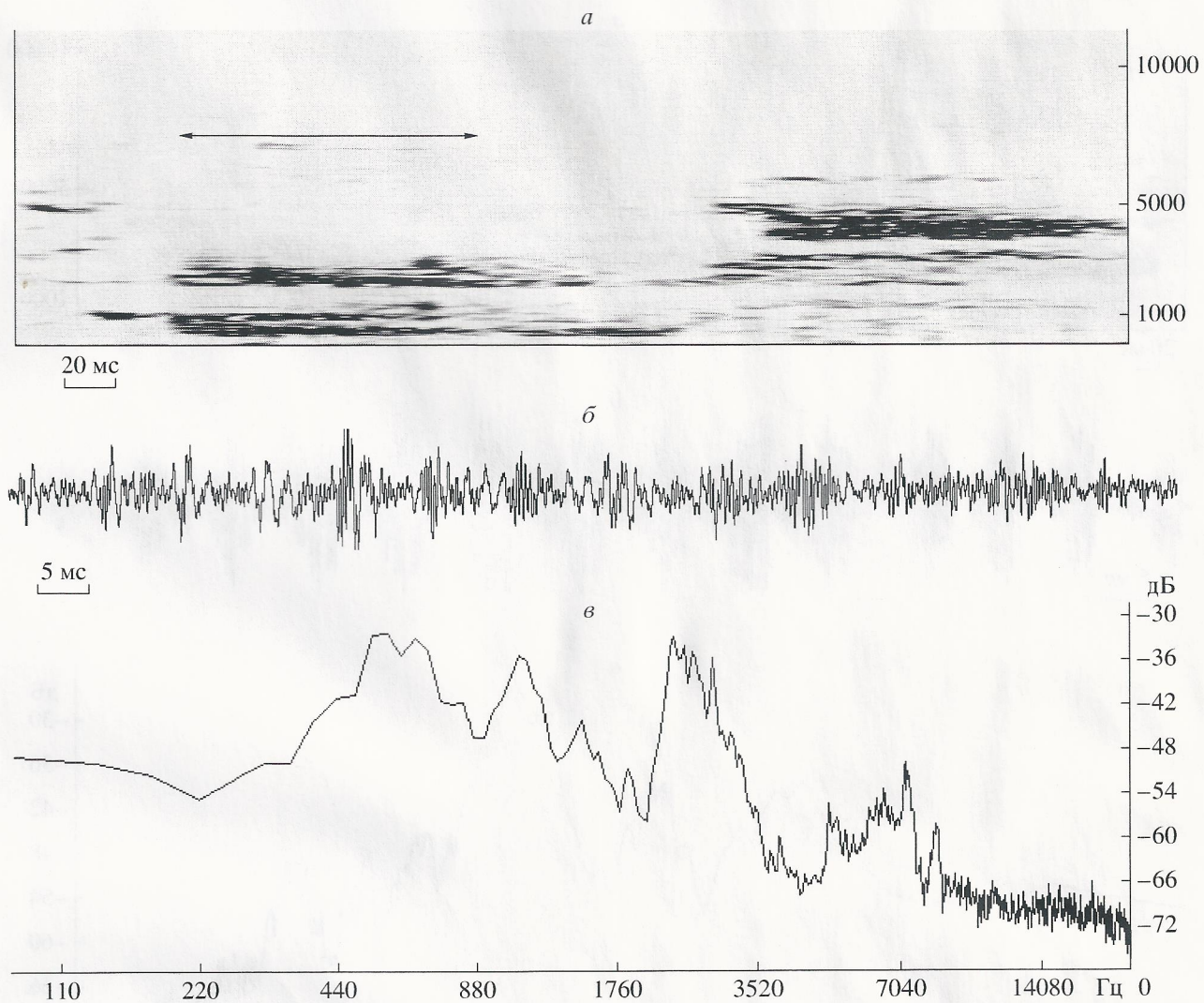


Рис. 2. Акустические характеристики звука "а" из слова "хороший", произнесенного серым попугаем Кузей. Обозначения, как на рис. 1.

различий распределений определяли по критерию Колмогорова-Смирнова.

РЕЗУЛЬТАТЫ

У серого попугая Кузи было выделено всего 786 гласноподобных участков. Из них к категории "а" было отнесено 303 звука, к "о" – 53 звука, к "у" – 102 звука, к "и" – 170 звуков (остальные звуки были отнесены к другим категориям или интерпретированы неоднозначно). У волнистого попугайчика Геши было выделено 968 гласноподобных звуков. К "а" отнесено 384 звука, к "о" – 59, к "у" – 38, к "и" – 163.

Акустические характеристики гласноподобных звуков

У Кузи и Геши звуки с гармонической структурой встречались редко. Формантные области у

волнистого попугайчика были выражены четче, чем у серого попугая. Частотный диапазон большей части произнесений у обеих птиц составлял 400–4000 Гц, причем для шипящих согласных он мог достигать 8 кГц. У Геши наиболее выраженные спектральные максимумы приходились на диапазон 2500–3500 Гц (рис. 1, а, в), у Кузи 1000–2500 Гц (рис. 2, а, в).

У волнистого попугайчика в спектре практически всех гласноподобных звуков присутствовал спектральный компонент с частотой от 2 до 4 кГц, по амплитуде равный или превышающий все остальные компоненты с более низкими частотами. Несущая частота гласноподобных звуков Геши, определенная на основании амплитудно-динамической формы (рис. 1, б), у 54% сигналов составляла 3–3.5 кГц и у 32% составляла 2.5 кГц. В оставшихся случаях она была ниже или изредка

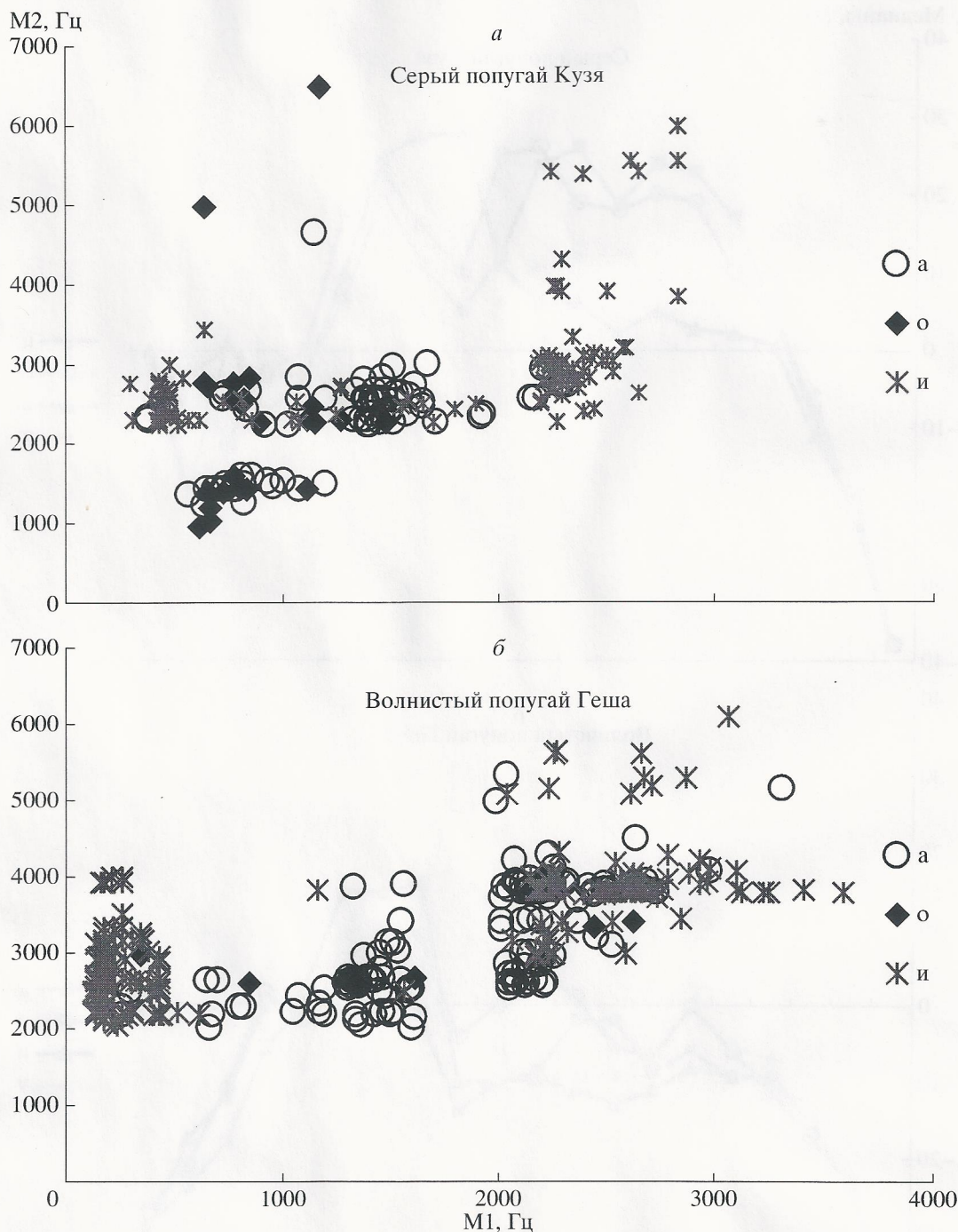


Рис. 3. Распределение значений частот двух наиболее выраженных спектральных компонентов для гласных звуков Кузи – а и Геша – б.

По оси абсцисс – частота первого из двух наиболее выраженных спектральных компонентов, Гц. По оси ординат – частота второго из двух наиболее выраженных спектральных компонентов, Гц.

вообще не определялась. Модулирующая частота была определена в 95% сигналов и составляла 90–140 Гц. У Кузи (рис. 2, б) несущая составляла 2–2.5 кГц в 41% сигналов и 3–3.5 кГц в 27%. В остальных случаях она была ниже или не определялась. Модулирующая частота колебалась от

100 до 500 Гц и могла быть определена только в 43 % случаев.

Аудиторский анализ

В данной работе был проведен аудиторский анализ четырехсот гласноподобных звуков попу-

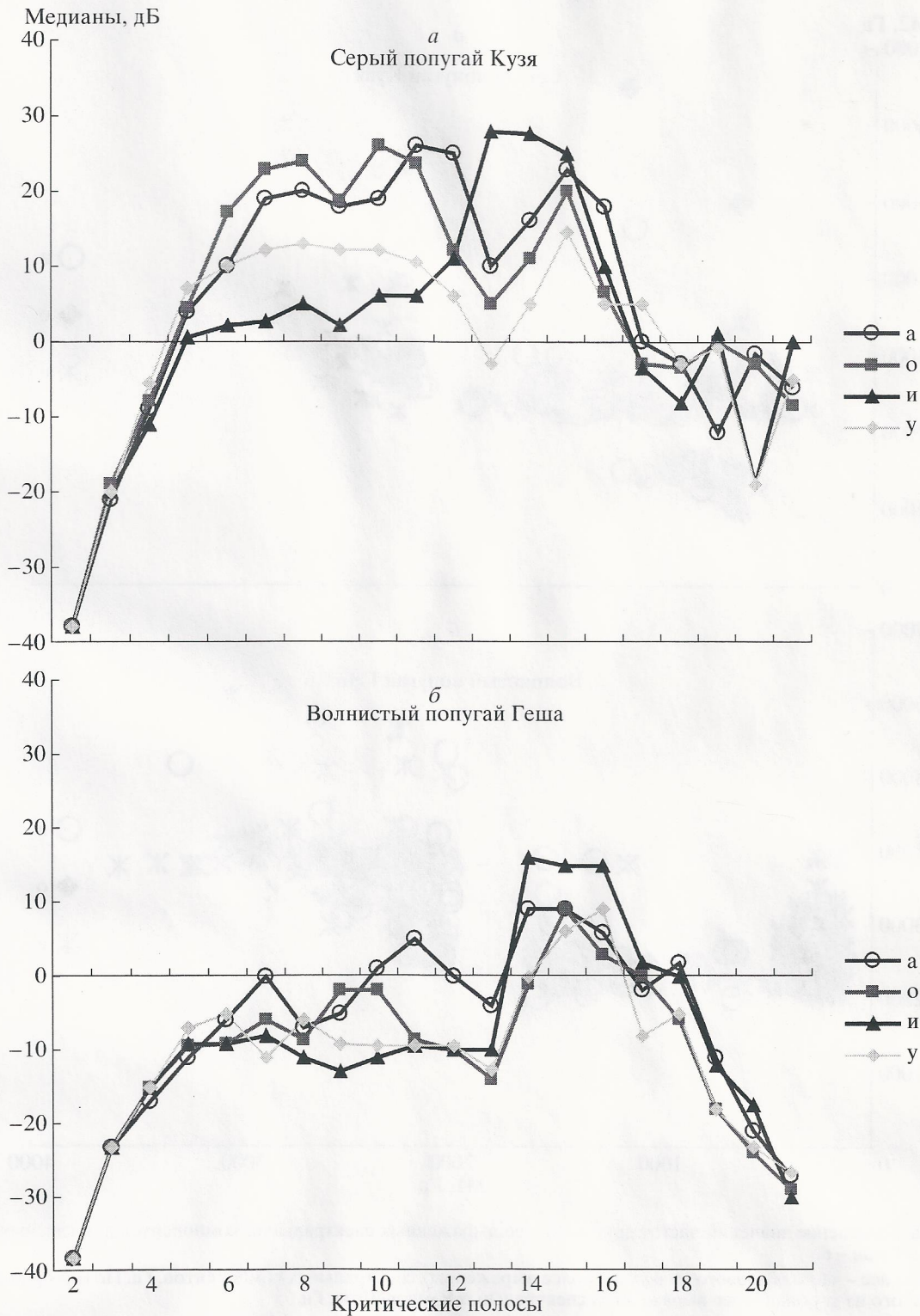


Рис. 4. Распределения зависимости медиан разности амплитуд от частоты спектральных компонентов для различных категорий гласных звуков, имитируемых попугаями Кузей – а и Гешей – б.

По оси абсцисс – номер критической полосы. По оси ординат – медианы разности амплитуд, дБ.

Соответствие критических полос частотным диапазонам: номер полосы (частотный диапазон); 2 – (100–200 Гц); 3 – (200–300 Гц); 4 – (300–400 Гц); 5 – (400–510 Гц); 6 – (510–630 Гц); 7 – (630–770 Гц); 8 – (770–915 Гц); 9 – (915–1080 Гц); 10 – (1080–1260 Гц); 11 – (1260–1480 Гц); 12 – (1480–1720 Гц); 13 – (1720–1990 Гц); 14 – (1990–2310 Гц); 15 – (2310–2690 Гц); 16 – (2690–3125 Гц); 17 – (3125–3675 Гц); 18 – (3675–4350 Гц); 19 – (4350–5250 Гц); 20 – (5250–6350 Гц); 21 – (6350–10000 Гц).

Достоверность различия медиан для пар звуков попугаев Геши и Кузи

Кузя

Пары \ №	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
а/о						+	+		+		+	+		+						
а/у						+	+	+	+	+	+	+	+	+	+					
а/и			+	+	+	+	+	+	+	+	+		+	+	+	+				
о/у						+	+	+	+	+	+		+							
о/и					+	+	+	+	+	+	+		+	+	+	+				
и/у			+	+	+	+	+	+	+			+	+	+	+					

Геша

Пары \ №	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
а/о						+		+		+	+	+	+			+	+	+		
а/у				+					+	+	+	+	+			+	+	+		
а/и				+		+	+	+	+	+	+	+	+	+	+	+			+	+
о/у																+				
о/и								+	+				+	+	+	+		+	+	
и/у													+		+	+	+	+		

В строках указаны пары сравниваемых звуков, в столбцах – номера критических полос, в которых проводилось сравнение. Знаком (+) указано значение $p \leq 0.05$ (различие достоверно).

гая Кузи и пятисот попугая Геши. Достоверно идентифицированы у Кузи были 78 “а”, 24 “о”, 34 “у”, 108 “и”; 73 звука были достоверно отнесены к двум категориям, шесть – к трем. Чаще всего встречались пары “а/о” (45% случаев), “а/э” и “о/у” (по 15%).

У Геши были определены 135 звуков “а”, 10 – “о”, 194 – “и”; 85 звуков были отнесены к двум категориям, шесть – к трем. Из них 28% пришлось на пары а/и и 22% на а/о.

Анализ значений частот двух наиболее выраженных спектральных компонентов

Распределения значений частот двух наиболее выраженных спектральных максимумов для достоверно идентифицированных звуков обоих попугаев на двухформантной плоскости точно так же, как и для всей массы звуков, перекрывались (рис. 3, а, б).

Анализ распределений зависимости медиан разности амплитуд от частоты спектральных компонентов

Построение распределений зависимости разности амплитуд спектральных компонентов от их частоты для идентифицированных гласноподобных звуков серого попугая (рис. 4, а) позволило разделить категории “о” и “у”, при этом распределения для “а” и “о” в значительной степени сблизились по сравнению с распределениями, полученными для всех выделенных звуков. Звук “и”

обладал одним максимумом спектра в высокочастотной области в 14–16 критических полосах. Максимумы спектра звука “у” находились в том же частотном диапазоне, что и для “о”, но значения разности амплитуд у “у” в 6–12 полосах оказались значительно меньше. У звука “а” по сравнению со звуком “о” в 10–13 критических полосах имелись спектральные компоненты, сдвинутые на одну критическую полосу в более высокочастотную область. Это отличие было ранее показано и для других птиц.

У волнистого попугайчика Геши частоты наиболее выраженных спектральных компонентов оказались примерно на 1 кГц выше, чем у более крупных птиц; разность амплитуд была крайне низкой, а наиболее выраженные спектральные максимумы находились в 14–16 полосах для всех категорий гласных. Так как не было обнаружено принципиальных отличий распределения зависимости медиан и разности амплитуд спектральных компонентов между фонетическими категориями, выделенными исследователем и идентифицированными аудитором, то анализировались распределения, построенные на основании данных по звукам, выделенным исследователем (рис. 4, б). Второй по величине разности амплитуд максимум для категории “а” приходился на 10–13 критическую полосу. Для звука “о” второй максимум совпал с 9–11 полосой, при этом разность амплитуд у него оказалась ниже, чем для категории “а”.

Для категории “у” распределение практически полностью совпало с распределением значений для категории “и”, у которой единственный максимум находился в 14–16 критических полосах.

Результаты статистической проверки достоверности отличий медиан разности амплитуд, по критерию Колмогорова-Смирнова, для Кузи и Геши представлены в таблице.

ОБСУЖДЕНИЕ

Результаты данной работы позволяют предположить, что амплитудная модуляция, показанная для волнистых попугайчиков (Banta Lavenex, 1999), также может иметь место и для серых попугаев. Но при этом у серых попугаев чаще встречается непериодическая амплитудная модуляция, а у волнистых – периодическая.

Наличие наиболее выраженного спектрального компонента в диапазоне 3–3.5 кГц у волнистого попугайчика совпала с несущей частотой, определенной по амплитудно-динамической форме сигнала, и попала в диапазон видовых вокализаций птиц, что и могло обуславливать его высокую амплитуду. Модулирующая частота в диапазоне 90–140 Гц совпадала с диапазоном частоты основного тона человека, обучавшего птицу, поэтому имитации Геши на слух звучали как низкочастотные. Можно предположить, что образование звуков путем амплитудной модуляции позволяет генерировать низкие частоты при том, что основной источник колебаний физически на это не способен.

В результатах аудиторского анализа обращает на себя внимание двойная идентификация звуков как “а” и “и” почти в 30% случаев. Возможно, что существуют индивидуальные отличия в восприятии различных категорий гласных. При наличии в звуках “а” наиболее выраженного спектрального компонента с частотой 3–3.5 кГц они воспринимались как “и”, так как наличие высокочастотных максимумов в спектре характерно как раз для этой категории звуков. У Кузи высокочастотный компонент чаще всего был ниже по амплитуде, чем остальные компоненты, и подобной двойной идентификации у него не встречалось. У обеих птиц лучше всего идентифицировались звуки “а” и “и”. Звуки “о” и “у” идентифицировались меньше чем в 30% случаев.

Для большинства пар звуков у Кузи и Геши достоверные отличия медиан разницы амплитуд спектральных компонентов имели место в значительном количестве полос. Сравнивая данные по этим попугаям с ранее полученными для серого попугая Ромы и майны Чики, можно отметить, что только в диапазонах 11–14 критических полос для пары категорий “а” и “о” достоверные отличия

имели место у всех четырех птиц. Для идентификации категории “и”, скорее всего, играет главную роль спектральный компонент в 15–17 критических полосах. Причем для волнистого попугайчика он совпадает с наиболее выраженным спектральным компонентом, соответствующим несущей частоте вокализаций, и попадает в частотный диапазон видоспецифических сигналов.

ВЫВОД

У волнистого и серого попугаев зависимость относительных амплитуд спектральных компонентов от их частоты для категорий гласных “а”, “о”, “у”, “и” имеет различный характер.

Работа выполнена при поддержке ведомственной научной программы Минобразования и науки России “Развитие научного потенциала высшей школы”, № 4727.

СПИСОК ЛИТЕРАТУРЫ

- Куликов Г.А., Андреева Н.Г., Павликова М.И., Самокищук А.П. Характеристики гласноподобных звуков детей первого полугодия жизни // Доклады Академии Наук. 1999. Т. 368. № 6. С. 843–845.
- Уплисова К.О. Характеристики гласных звуков, имитируемых говорящими птицами // Сенсорные системы. 2004. Т. 18. № 3. С. 199–205.
- Цвикер Э., Фельдкеллер Р. Ухо как приемник информации / Пер. с нем. М.: Связь, 1965. 104 с.
- Banta Lavenex P. Vocal production mechanisms in the budgerigar: The presence and implications of amplitude modulation // J. Acoust. Soc. Amer. 1999. V. 106. № 1. P. 491–505.
- Dooling R.J., Best C.T., Brown S.D. Discrimination of synthetic full-formant and sinewave /ra-/la/ continua by budgerigars and zebra finches // J. Acoust. Soc. Amer. 1995. V. 97. № 3. P. 1839–1846.
- Gaunt A.S., Gaunt S.L.L. Electromyographic studies of the syrinx in parrot (Aves, Psittacidae) // Zoomorphol. 1985. V. 105. P. 1–11.
- Klatt D.H., Stefanski R.A. How does a mynah bird imitate human speech? // J. Acoust. Soc. Amer. 1974. V. 55. № 4. P. 822–832.
- Patterson D.K., Pepperberg I.M. A comparative study of human and parrot phonation: Acoustic and articulatory correlates of vowels // J. Acoust. Soc. Amer. 1994. V. 96. № 2. P. 634–648.

The Acoustic and Auditor Analysis of Vowel-like Sounds of Gray Parrot (*Psittacus erithacus*) and Budgerigar (*Melopsittacus undulatus*)

K. O. Uplisova

The St.-Petersburg State University 199034, St.-Petersburg, University emb., 7/9

In connection with studying of acoustic characteristics of vowel sounds under different conditions of production the acoustic analysis of vowel-like sounds of grey parrot and budgerigar was realized. As it has already been proved (Patterson, Pepperberg, 1994), that vowel-like sounds of talking birds cannot be divided into categories on the basis of frequencies of formants (spectral maxima) that is characteristic for human speech, in the given work the attribute of dependence of relations of amplitudes of spectral components from their frequency, applied for vowel-like sounds of babies (Kulikov et al., 1999) has been used. Results of the analysis have shown, that dependence of relations of amplitudes on frequency of spectral components of phonetic categories "a", "o", "u", "i" for grey parrot and budgerigar are different.

Key words: the imitation, talking birds, a parrot, vowel sounds.