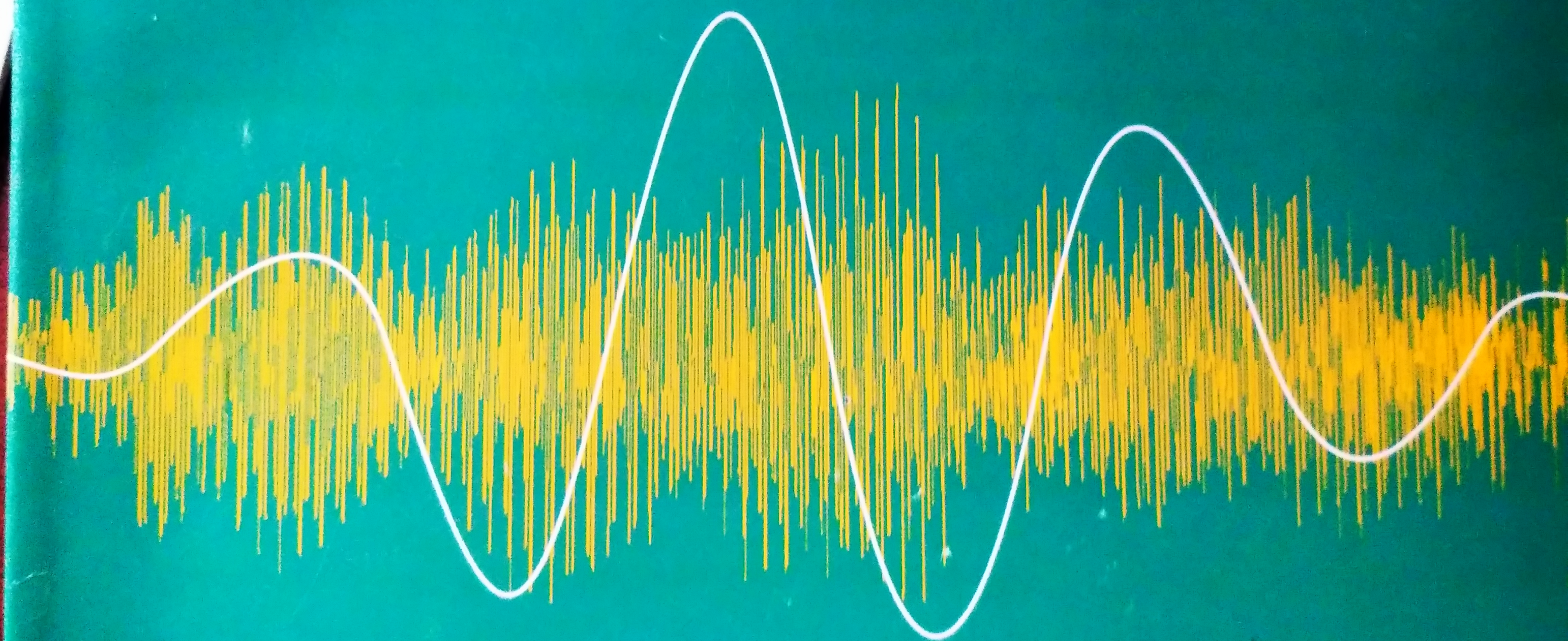


СБОРНИК ТРУДОВ  
**XX СЕССИИ**  
РОССИЙСКОГО  
АКУСТИЧЕСКОГО  
ОБЩЕСТВА



акустика речи  
медицинская и биологическая акустика  
архитектурная и строительная акустика  
шумы и вибрации  
аэроакустика

**ТОМ 3**

2018-12-11 14:02



Российская академия наук  
Российская академия естественных наук  
Российское акустическое общество  
Акустический институт имени академика Н.Н.Андреева  
Московский государственный горный университет

# **СБОРНИК ТРУДОВ**

**XX СЕССИЯ  
РОССИЙСКОГО АКУСТИЧЕСКОГО ОБЩЕСТВА**

27-31 октября 2008 года

**Акустика речи  
Медицинская и биологическая акустика  
Архитектурная и строительная акустика  
Шумы и вибрации  
Аэроакустика**

Москва  
ГЕОС  
2008

2018-12-11 14:02



УДК 534

ISBN 978-5-89118-429-9

Акустика речи. Медицинская и биологическая акустика. Архитектурная строительная акустика. Шумы и вибрации. Аэроакустика. Сборник трудов XX сессии Российского акустического общества. Т. 3. - М.: ГЕОС, 2008. 328 с.

В книге собраны доклады XX сессии Российского акустического общества

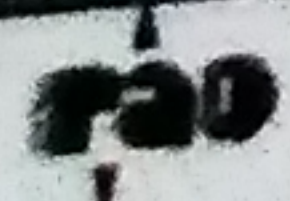
### ОРГАНИЗАЦИОННЫЙ КОМИТЕТ XX сессии Российского акустического общества

- Руденко Олег Владимирович – Председатель оргкомитета, академик РАН, зав.каф. акустики МГУ  
Корчак Андрей Владимирович – Зам. Председателя оргкомитета, профессор, ректор МГТУ  
Акуличев Виктор Анатольевич – Зам. Председателя оргкомитета, Президент РАО, академик РАН, директор ТОИ ДВО РАН  
Гурбатов Сергей Николаевич – Зам. Председателя оргкомитета, профессор, проректор ННГУ  
Бобровницкий Юрий Иванович – профессор, зав. отделом ИМАШ РАН  
Кедринский Валерий Кириллович – профессор, зам. директора ИГЛ СО РАН  
Маслов Вячеслав Львович – профессор, зам. директора ФГУП «ЦНИИ им. акад. А.Н.Крылова»  
Осипов Георгий Львович – директор НИИ СФ РААСХ  
Рубан Анатолий Дмитриевич – член-корр., зам дир. Ин-та компл. освоения недр РАН  
Тимошенко Владимир Иванович – академик РАЕН, профессор ТГРТУ  
Шкуратник Владимир Лазаревич – профессор, зав. каф. МГГУ  
Юдина Елена Васильевна – исп. директор Российского акустического общества

### ПРОГРАММНЫЙ КОМИТЕТ

- |   |  |
|---|--|
| Абрамов Олег Владимирович – ИОХ РАН           | Кузнецов Владимир Михайлович – ЦАГИ        |
| Бибиков Николай Григорьевич – АКИН            | Лопашев Дмитрий Захарович – ВНИИФТРИ       |
| Борисов Лев Александрович – НИИ СФ РААСХ      | Малеханов Александр Игоревич – ИПФ РАН     |
| Вознесенский Александр Сергеевич – МГГУ       | Миронов Михаил Арсеньевич – АКИН           |
| Голямина Ирина Петровна – АКИН                | Петников Валерий Георгиевич – НЦВИ ОИФ РАН |
| Диденкулов Игорь Николаевич – ИПФ РАН         | Потапова Родмонга Кондратьевна – МГЛУ      |
| Егерев Сергей Викторович – АКИН               | Прончатов-Рубцов Николай Васильевич – ННГУ |
| Ерофеев Владимир Иванович – Ниж.фил. ИМАШ РАН | Рыбак Самуил Акивович – АКИН               |
| Есипов Игорь Борисович – АКИН                 | Сорокин Виктор Николаевич – ИППИ РАН       |
| Копьев Виктор Феликсович – ЦАГИ               | Фокин Андрей Викторович – АКИН             |
| Куличков Сергей Николаевич – ИФА РАН          |  |

Труды сессии изданы без дополнительного редактирования по оригиналам документов представленных авторами. Оргкомитет не несет ответственности за стиль и фактологическую изложенные представленных материалов.



Российское акустическое общество

Российский фонд фундаментальных исследований (08-02-06168)

2018-12-11 14:03



## XX сессия Российского акустического общества

2. Дмитриев Л. Б., Чернов Б. П., Маслов В. Т. Тайна тувинского «дуэта» или свойство гортани человека формировать механизм аэродинамического свиста. Новосибирск, 1992.
3. Ондар М. А. Х., Сарыглар А. С. О физической природе звуков тувинского горлового пения // Вопросы изучения истории культуры народов Центральной Азии и сопредельных регионов: материалы Междунар. науч.-практ. конф., г. Кызыл, 5-8 сентября 2005 г. Кызыл, 2006. С.380-381.
4. Гоноровский И. С. Радиотехнические цепи и сигналы. Часть II. М.: Советское Радио, 1967.

УДК 591.582.2

К.О. Уплисова

## ГОВОРЯЩИЕ ПТИЦЫ: ОСОБЕННОСТИ МОДИФИЦИРОВАННЫХ ГЛАСНЫХ ЗВУКОВ

Санкт-Петербургский государственный университет  
Россия, 199034, Санкт-Петербург, Университетская наб., 7/9  
Тел.: (812) 328-97-06; E-mail: sehmet@fromgu.com

*Способность некоторых видов птиц имитировать человеческую речь представляет особый интерес, и в частности, потому что строение их звукогенерирующего аппарата существенно отличается от человеческого. Это приводит к тому, что резонансные характеристики голосового аппарата птиц не могут быть такими же, как у гортани человека. В результате в большинстве случаев значения формант гласных звуков говорящих птиц не соответствуют человеческим, тем не менее, гласные идентифицируются с высокой вероятностью. С целью выявления акустических признаков, позволяющих различать имитационные звуки, была проведена модификация серии гласных. Модификация проводилась путем подавления амплитуды спектральных компонентов, или транспонированием всех частот спектра. Было выявлено, что если в спектре звука присутствуют выраженные спектральные компоненты в диапазоне 1300–1700 Гц, этот звук будет с высокой вероятностью идентифицирован как «а», независимо от того присутствуют ли выраженные спектральные компоненты на низких и высоких частотах. Если частоты основных энергетически выраженных компонентов находятся ниже 1000 Гц, то звук будет идентифицирован как «о», независимо от присутствия выраженных спектральных компонентов выше 1700 Гц, но не в случае их наличия в диапазоне 1300–1700 Гц. В случае наличия энергетически выраженных спектральных компонентов с частотой ниже 500 Гц и при отсутствии выраженных спектральных компонентов на более высоких частотах звук будет идентифицироваться как «у». Если выраженные спектральные компоненты будут присутствовать выше 1700 Гц звук будет идентифицироваться как «и».*

**ВВЕДЕНИЕ** Актуальность проблемы распознавания и восприятия речи человеком отражена в большом количестве работ, посвященных разработке этих вопросов. Однако, несмотря на длительное исследование, до сих пор многие аспекты восприятия и распознавания речи остаются невыясненными, например, какие акустические характеристики звуков являются перцептивно-значимыми для идентификации фонетических категорий гласных. Согласно классическим представлениям [1, 2] ключевыми признаками гласных фонем являются частоты первой, второй и, иногда, третьей формант, отражающих резонансные частоты голосового тракта человека. Для каждой категории гласного значения частот первой и второй формант являются стабильными, и их значения не пересекаются со значениями формант звуков другой категории [3]. Однако, например, форманты звуков с высокой частотой основного тона (выше 300 Гц), присутствующих в гласноподобных вокализациях младенцев и звуках вокальной речи, не попадают в границы, определенные для звуков речи взрослого диктора. Тем не менее, подобные гласные идентифицируются аудитором с высокой вероятностью [4, 5].

В связи с этим, представляет особый интерес способность некоторых видов птиц имитировать человеческую речь. Ведь для этого птице необходимо, несмотря на индивидуальные различия голосов окружающих её людей и характеристики своей слуховой системы, настроенной на особенности видоспецифических сигналов, выделить те акустические признаки, на основании которых она создаст копию звука, слова или фразы. Строение звукогенерирующего аппарата птиц также существенно отличается от человеческого [6, 7] и, следовательно, его резонансные характеристики не могут быть такими же как у человека. Более того, некоторые виды птиц используют при звукогенерации два источника звуковых колебаний. Так как звуки речи, имитируемые птицами, адекватно распознаются человеком, естественно предположить, что в них имеются акустические признаки, позволяющие отнести звук к той или иной категории. [8]. В ранее проведенной работе [9, 10] были определены особенности спектра гласных звуков, стабильные для категорий гласных звуков «а», «о», «у» и «и», имитируемых разными птицами. Учет значений частоты и амплитуды всех спектральных компонентов фонем позволил выявить различия между звуками, относящимися к категориям «а», «о», «у», «и», а именно присутствие энергетически выраженных спектральных компонентов у звуков категории «о» в диапазоне



630-1260 Гц, «а» - 1260-1990 Гц, «у» - ниже 630 Гц, «и» - выше 1990 Гц. Однако для проверки предположения о том, что максимумы, находящиеся именно на этих частотах, являются критическими для идентификации той или иной категории было необходимо подвергнуть звуки модификации и выявить изменилась ли при этом фонетическая идентификация звука.

Таким образом данная работа посвящена исследованию влияния изменения акустических характеристик спектральных максимумов на фонетическую идентификацию гласных звуков.

**МЕТОДИКА** Для модификации спектральных компонентов были отобраны гласные звуки серого и волнистого попугаев с высокой вероятностью идентифицировавшиеся как «а», и «о». С помощью фильтра преобразования Фурье (FFT фильтр, программа Cool Edit Pro) производились изменения амплитуды спектральных составляющих. Если в спектре звука присутствовали выраженные спектральные максимумы, которые можно было характеризовать единственным значением частоты и амплитуды, то проводилось понижение их амплитуды таким образом, чтобы этот максимум не мог бы быть выделен (первая серия модификаций). В случае если спектр звука характеризовался выраженной полосой спектральных компонентов, проводилось изменение его ширины путем подавления амплитуды спектральных составляющих в его определенной части (вторая серия модификаций). Для транспонирования частот (третья серия) использовалась функция Constant Stretch программы Cool Edit Pro. Транспонирование в сторону понижения частоты осуществлялось шагами с отношением 0,79 всех частот предыдущего звука по отношению к частотам последующего. То есть если у исходного звука значения частот формант составляли 600 Гц, 1270 Гц, 2150 Гц, то в результате первого шага транспонирования в сторону понижения, частоты формант составили 474 Гц, 1003 Гц, 1698 Гц, в результате второго шага – 374 Гц, 792 Гц и 1341 Гц и т.д. Повышение частот осуществлялось с соблюдением отношения 1,33. В этом случае частоты формант в результате первого шага в сторону повышения частоты составили 798 Гц, 1689 Гц, 2859 Гц. Для каждого звука было проведено несколько шагов транспонирования частоты в сторону ее понижения и повышения. Модифицированные звуки прослушивались группой из семи аудиторов, постоянно участвующих в подобных экспериментах. Перед ними ставилась задача обозначить изменил ли модифицированный звук категорию по сравнению с категорией не модифицированного звука и каким образом.

**РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ** Первая модификация заключалась в подавлении амплитуды спектральных компонентов ниже 750 Гц и выше 1700 Гц. В этом случае для всех звуков категорий «а» и «о» изменение идентификации было зафиксировано не более чем у 3-х из семи аудиторов, что позволило предположить, что критические признаки категорий «а» и «о» находятся в диапазоне 750-1700 Гц.

В первой серии модификаций было модифицировано четыре звука волнистого попугайчика (три «а» и один «о») и три звука серого попугая (два «а» и один «о»). В трех двухформантных звуках волнистого попугайчика элиминация первого максимума, а в одном трехформантном, элиминация первого и второго максимума, привела к существенному понижению идентификации исходных звуков как «а» и «о». Модифицированные двухформантные звуки «а» и «о» стали идентифицироваться как «и» всеми аудиторами, оценка категории модифицированного трехформантного «а» также изменилась у шести аудиторов, но менее единообразно (рис. 1, 1). Однако при сохранении первых двух максимумов у трехформантного «а» (рис. 1, 2) существенного изменения идентификации категории выявлено не было. Не повлияло на оценку категории «а» и подавление первой форманты (470–860 Гц) (рис. 1, 3), в то время как подавление второй (1000–1700 Гц) привело к изменению оценки категории практически всеми аудиторами (рис. 1, 4). Подавление второй и третьей форманты изменило оценку категории «а» на категорию «о» большинством аудиторов (рис. 1, 5). У двух модифицированных трехформантных «а» серого попугая элиминация второй форманты (1000–1700 Гц) вызвала идентификацию звука как «о» не менее чем пятью аудиторами. Элиминация первой (400–900 Гц) не повлияла на идентификацию звука как «а». Таким образом для звуков категорий «а» подавление амплитуды форманты в области 900–1700 Гц приводит к изменению идентификации категории исходного звука «а». В этом случае чаще всего звук начинает идентифицироваться как «о».

Во второй серии было модифицировано четыре звука «а» серого попугая. У звука «а» с энергетически выраженной полосой 700–1600 Гц подавление амплитуды спектральных компо-

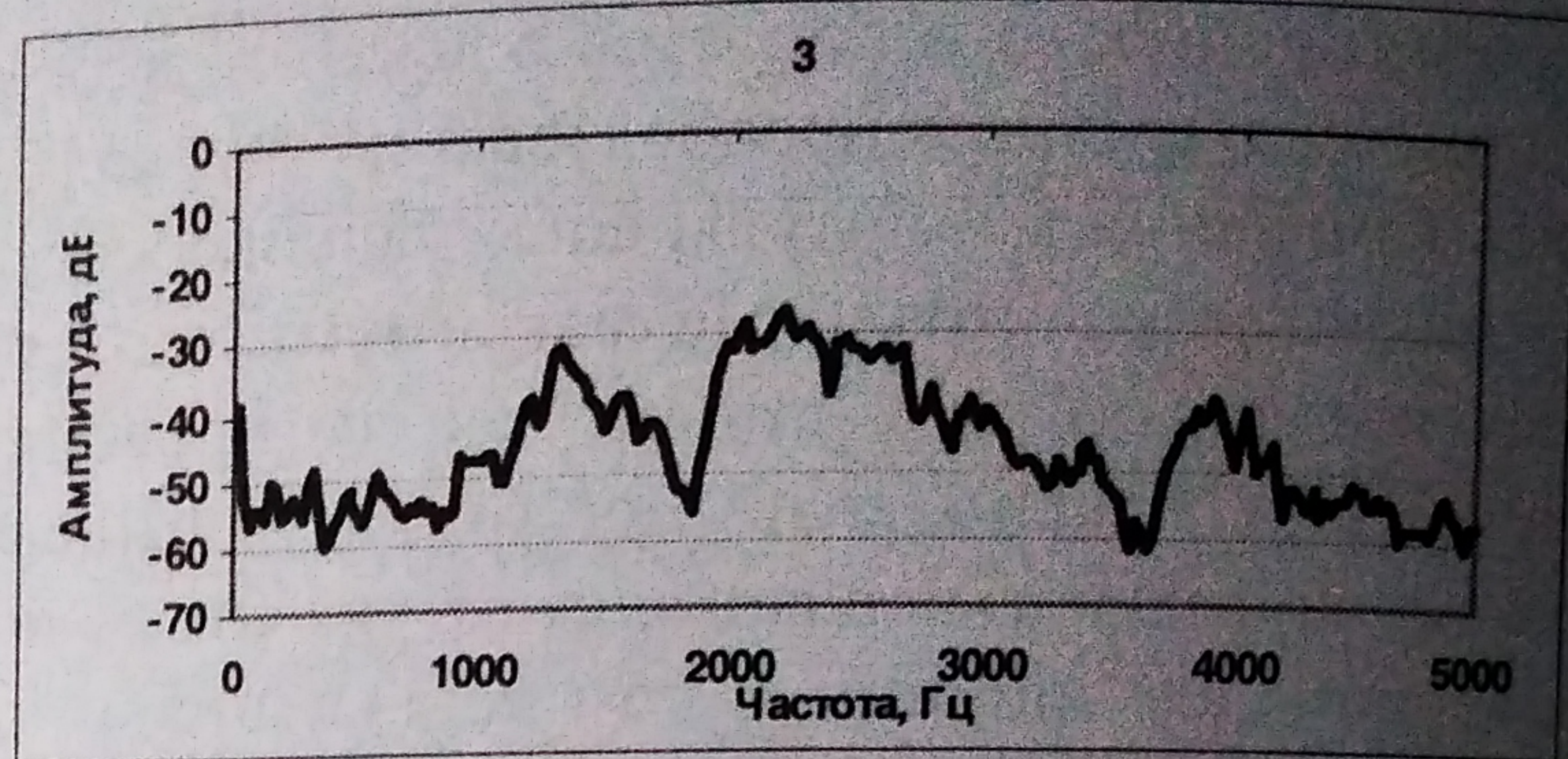


XX сессия Российского акустического общества

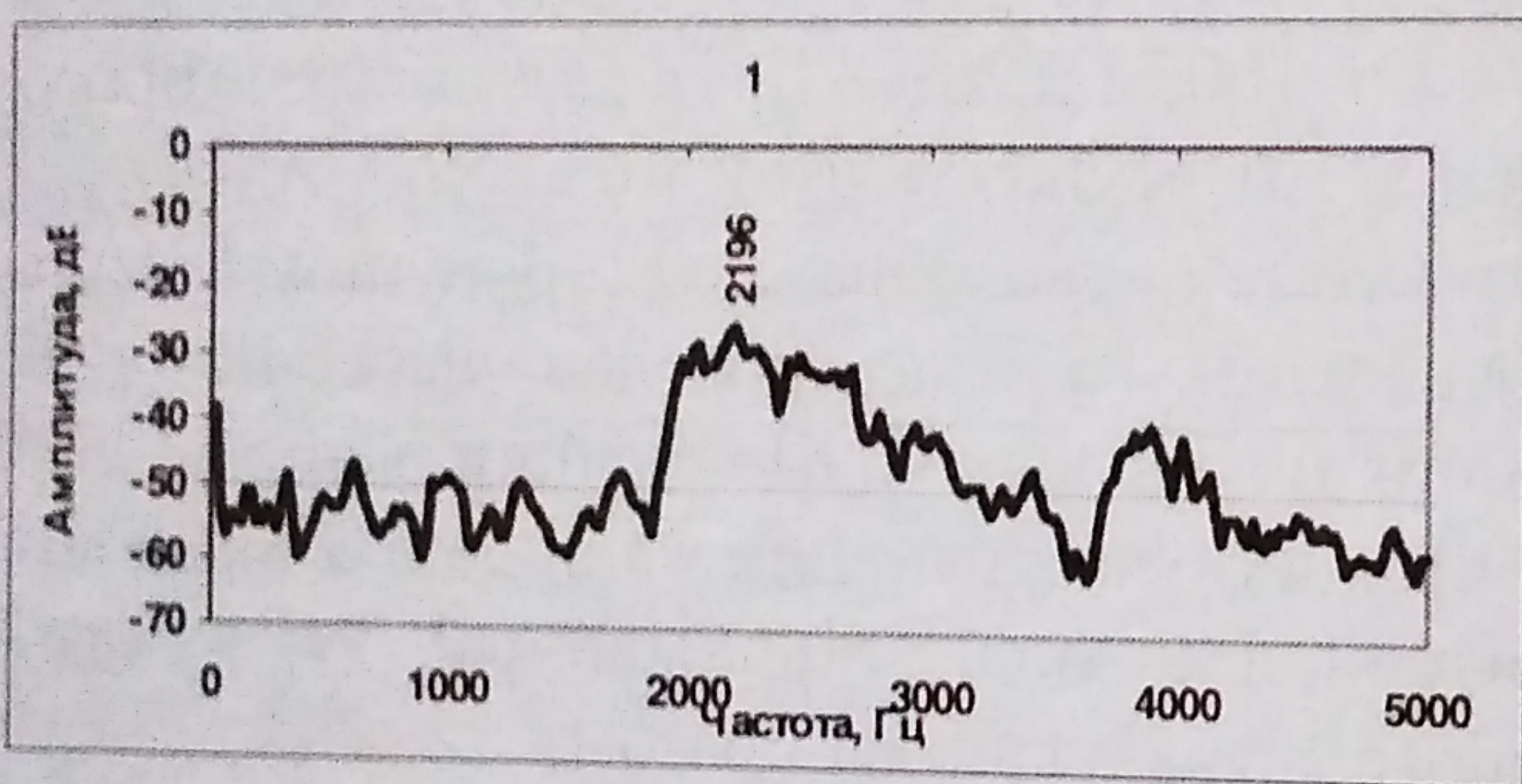
нентов в частотном диапазоне 1300–2000 Гц не вызвало изменения идентификации категории. Подавление амплитуды спектральных компонентов в диапазоне 1000–2000 Гц привело к идентификации звука как «о» всеми аудиторами, подавление амплитуды в диапазоне 800–2000 Гц вызвало появление оценок звука как «у» частью аудиторов. Сходная картина наблюдалась и для остальных звуков. Это позволило прийти к выводу о том, что при наличии энергетически выраженных компонентов не выше 1000–1100 Гц звук будет идентифицироваться как «о», не выше 600–800 Гц – как «у».



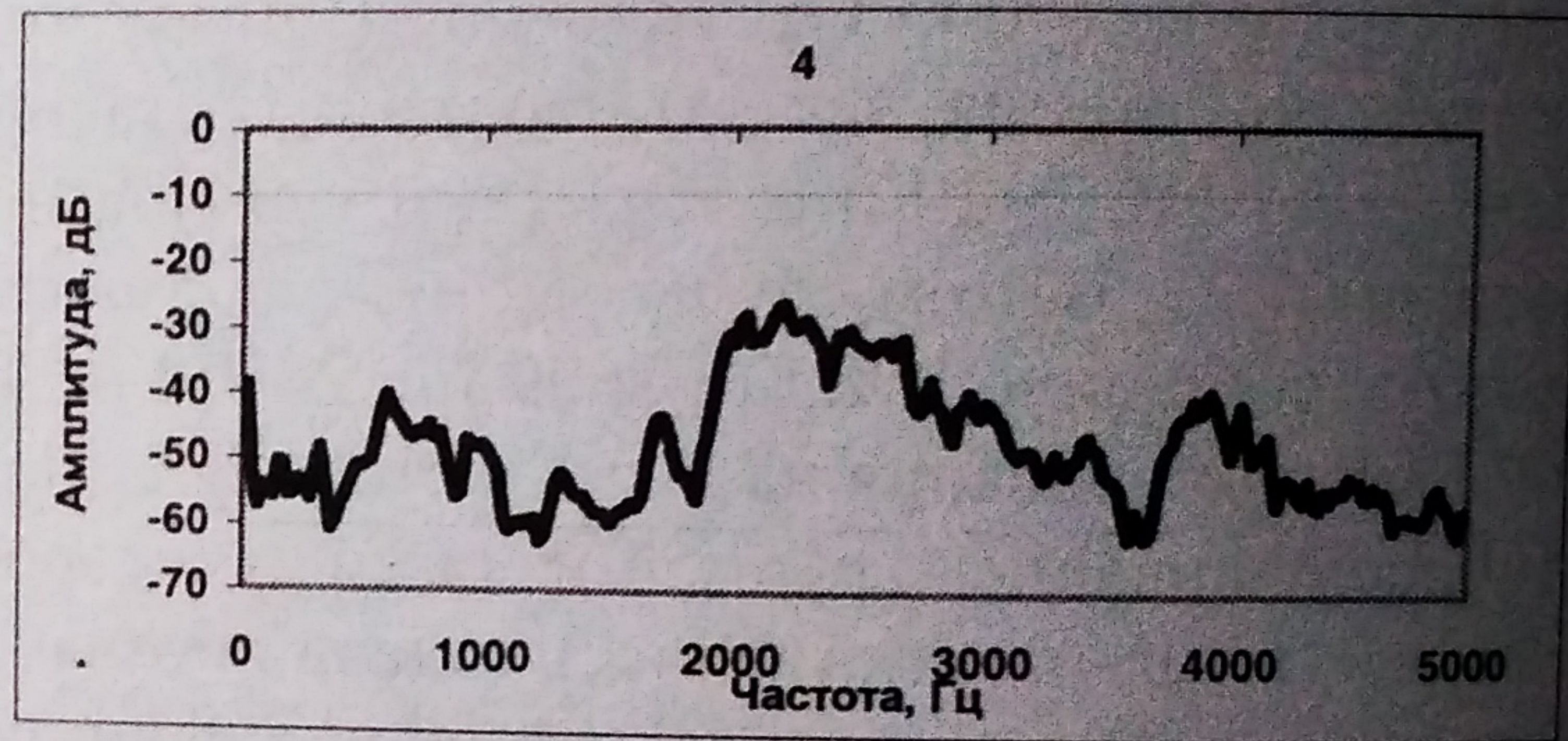
Исходный звук  
Не модифицированный звук  
Идентификация: 7 – «а»



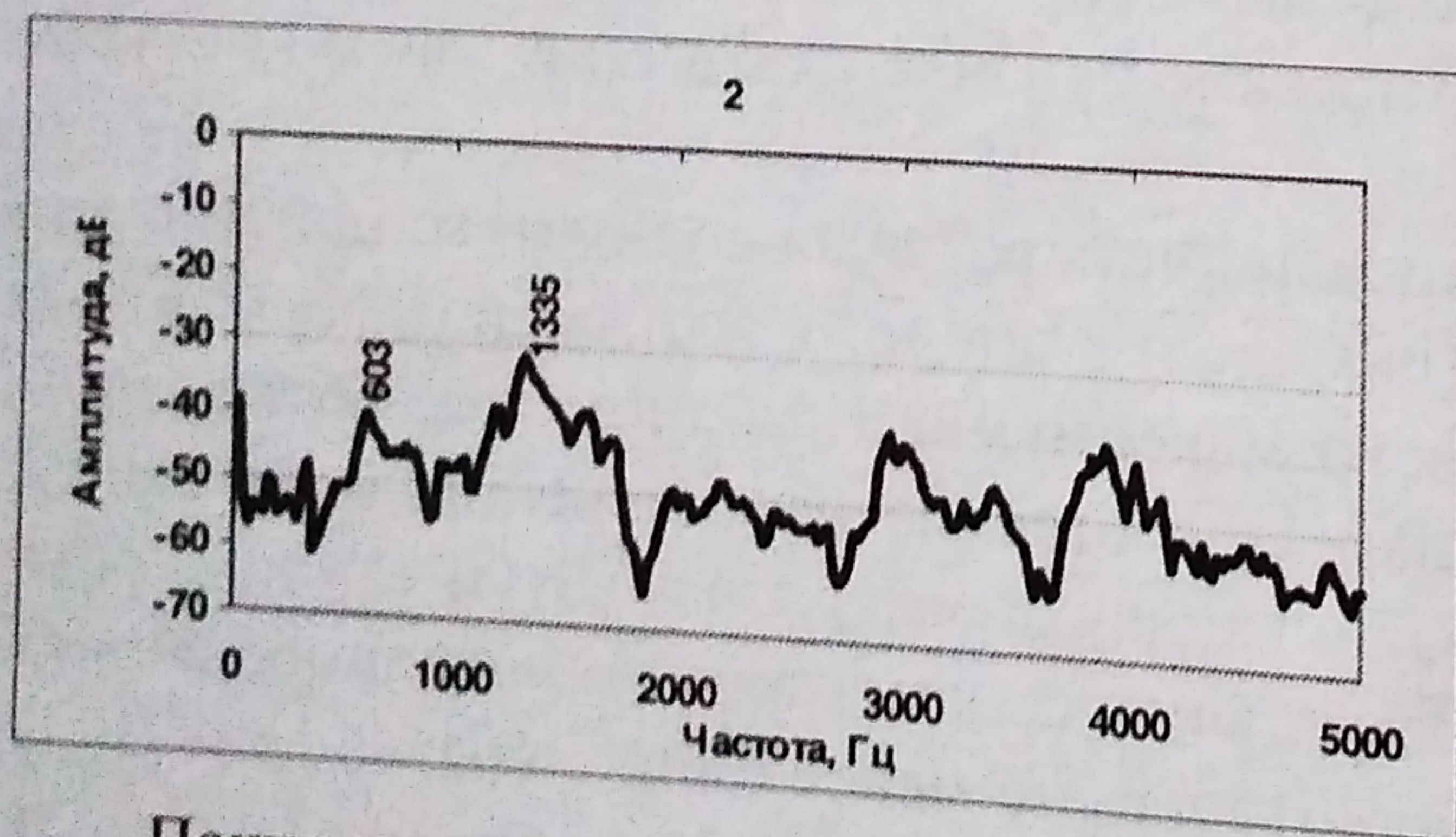
Понижение амплитуды 470-860 Гц  
Идентификация: 7 – «а».



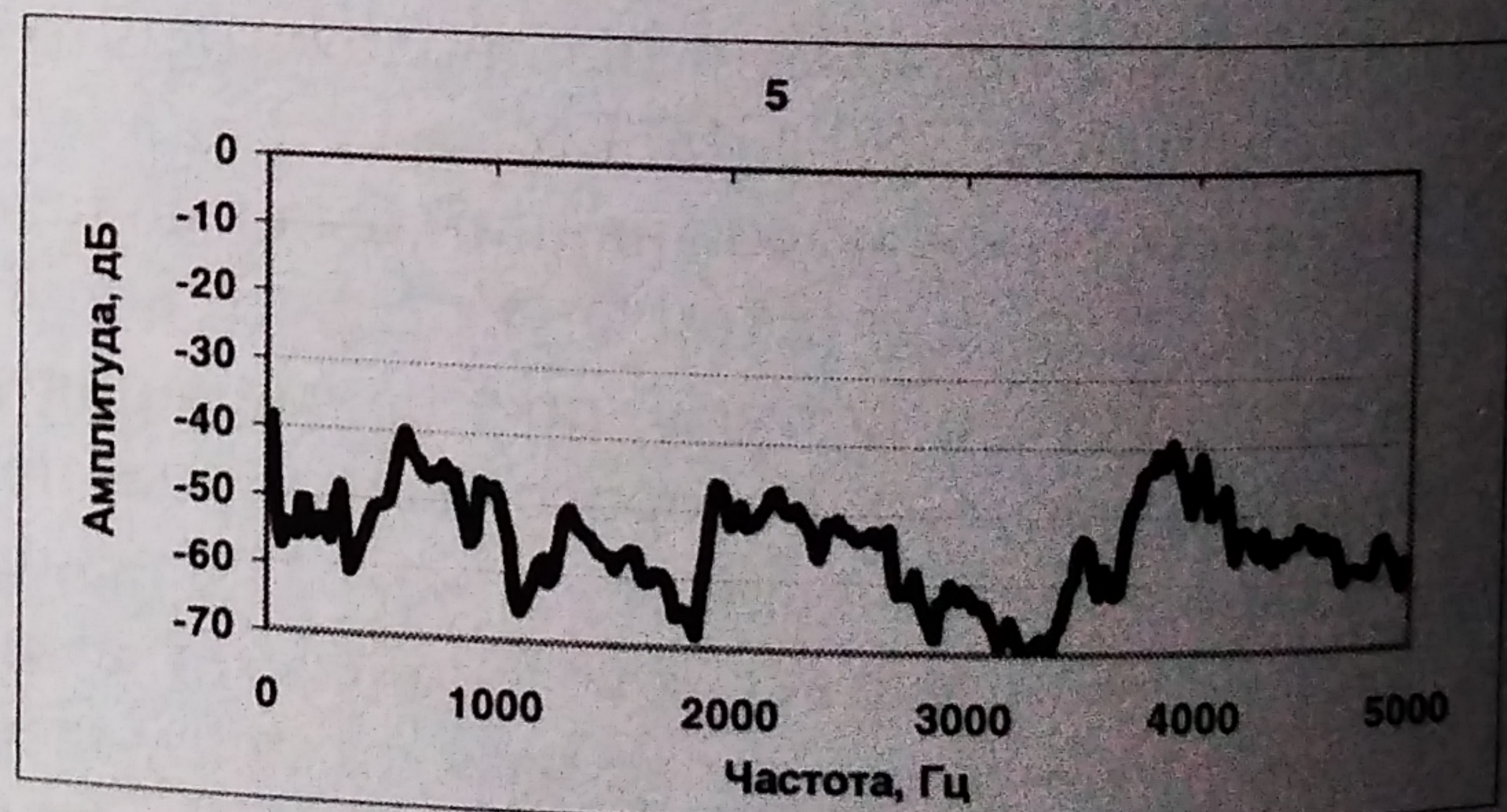
Понижение амплитуды 470-1765 Гц  
Идентификация: 3 – «и», 1 – «а», 1 – «о», 1 – «е», 1 – «э».



Понижение амплитуды 1000-1700 Гц  
Идентификация: 5 – «э», 1 – «е», 1 – «и».



Понижение амплитуды 1722-2900 Гц  
Идентификация: 4 – «а», 2 – «о», 1 – «у».



Понижение амплитуды 1000-3500 Гц  
Идентификация: 5 – «о», 2 – «у».

Рис. 1. Спектры модифицированных гласных «а» волнистого попугайчика. Под спектрами указан диапазон частот, в котором проводилось понижение амплитуды спектральных компонентов, и количество аудиторов давших ту или иную оценку

В третьей серии у двух «а» волнистого попугайчика и одного «а» серого попугая были транспонированы частоты. При транспонировании частот двух звуков «а» волнистого попугайчика значения частот формант исходных звуков составили: 600, 1270, 2150 Гц и 1380, 2600 Гц (двухформантный). В результате первого транспонирования в сторону понижения частоты формант составили 450, 990, 1636 и 1033, 1980 Гц. Уже первый шаг транспонирования привел к изменению идентификации звука частью аудиторов. Звук стал идентифицироваться как «о» ча-



стью аудиторов. Только к четвертому шагу для трехформантного звука (частоты формант 193, 409, 689 Гц) все аудиторы зафиксировали изменение категории «а» на «о» или «у». У двухформантной: три аудитора определили звук как «о» и «у», двое как «а» и «э». Повышение частотного звука. При модификации трехформантного звука и пять шагов для двухформантного идентифицировали звук как «и» (частоты формант 750, 1636, 2700 Гц). Все аудиторы определили категорию «и» при четвертом шаге (частоты формант 1530, 3337, 5770 Гц). Двухформантный «а» был идентифицирован как «и» всеми аудиторами при третьем шаге транспонирования в сторону повышения частоты.

При транспонировании частот в сторону понижения «а» серого попугая с энергетически выраженными полосами частот 750–1800 Гц и 2200–2800 Гц уже при первом шаге трое аудиторов идентифицировали звук как «о». Ко второму шагу двое из них идентифицировали звук как «у». К четвертому шагу трое аудиторов идентифицировали звук как «у», один как «о» и для одного аудитора оценка категории исходного звука не изменилась ни на одном из модифицированных звуков. При транспонировании частот в сторону повышения только к шестому шагу двое аудиторов идентифицировали звук как «и», для остальных оценка категории «а» не изменилась.

Таким образом, транспонирование частот звуков приводит к изменению оценки категории «а» на «о» и «у» при понижении частоты, и на категорию «и» при повышении частоты.

**ВЫВОДЫ** Если в спектре звука присутствуют выраженные спектральные компоненты в диапазоне 1300–1700 Гц, этот звук будет с высокой вероятностью идентифицирован как «а», независимо от того присутствуют ли выраженные спектральные компоненты на низких и высоких частотах. Если частоты основных энергетически выраженных компонентов находятся ниже 1000 Гц, то звук будет идентифицирован как «о», независимо от присутствия выраженных спектральных компонентов выше 1700 Гц, но не в случае их наличия в диапазоне 1300–1700 Гц. При наличии спектральных компонентов в диапазоне 1100–1300 Гц, звук может идентифицироваться и как «а» и как «о». В случае наличия энергетически выраженных спектральных компонентов с частотой ниже 500 Гц и при отсутствии выраженных спектральных компонентов на более высоких частотах звук будет идентифицироваться как «у». В случае одновременного присутствия выраженных спектральных компонентов с частотой ниже 500 Гц и выше 1700 Гц, звук будет идентифицироваться как «и», также он будет идентифицироваться, если выраженные спектральные компоненты будут присутствовать только начиная с 1700 Гц.

*Работа поддержана грантом РГНФ №07-06-00546а*

#### ЛИТЕРАТУРА

1. Фант Г. Анализ и синтез речи. – Новосибирск: Наука, 1970. – 166 с.
2. Фланаган Дж. Анализ, синтез и восприятие речи. – М.: Связь, 1968. – 396 с.
3. Peterson G.E., Barney H.L. Control methods used in a study of the vowels // *Journal of the Acoustical Society of America*. – 1952. – V. 24. – P. 175-184.
4. Куликов Г.А. и др. Характеристики гласноподобных звуков детей первого полугодия жизни // Доклады Академии Наук. – 1999. – Т. 368, № 6. – С. 843-845.
5. Андреева Н.Г., Куликов Г.А. Характеристики вокальных гласных при разной частоте основного тона // *Рос. Физиол. Журн. им. И.М. Сеченова*. – 2003. – Т. 89, №6. – С. 715-724.
6. Goller F., Larsen O.N. New perspectives on mechanisms of sound generation in songbirds // *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*. – 2002. – Vol. 188. – № 11-12. – P. 841-850.
7. Gaban-Lima R., Höfling E. Comparative anatomy of the syrinx in the tribe Arini (Aves: Psittacidae) // *Braz. J. Morphol. Sci.* – 2006. – V. 23. – № 3-4. – P. 501-512.
8. Patterson D.K., Pepperberg I.M. A comparative study of human and parrot phonation: Acoustic and articulatory correlates of vowels // *JASA* – 1994. – Vol. 96. – № 2. – P. 634-648.
9. Уплисова К.О. Характеристики гласных звуков имитируемых говорящими птицами // *Сенсорные системы*. – 2004. – Т.18, №3. – С. 199-205.
10. Уплисова К.О. Акустический и аудиторский анализ гласноподобных звуков серого (*Psittacus erithacus*) и волнистого (*Melopsittacus undulatus*) попугаев // *Сенсорные системы*. – 2006. – Т.20, №3. – С. 229-237.