

### Часть.15.1

## **Слуховое восприятие пространственных систем ч. 1**

Ирина Алдошина

Последнее десятилетие характеризуется бурным развитием систем пространственной звукопередачи (недаром 109-й конгресс AES в Лос-Анжелесе назывался Surrounded by Sound). Разработка и широкое использование таких систем в значительной степени изменило технологию звукозаписи, принципы проектирования систем звуковоспроизведения и пр. Дальнейшее их развитие требует новой, более глубокой информации о различных аспектах пространственного слуха, без получения которой невозможно решать такие глобальные задачи, как проблемы переноса пространственного звукового образа из первичного помещения (концертного зала, студии и др.) во вторичное помещение прослушивания.

В статье "Бинауральный слух и пространственная локализация" ("Звукорежиссер" 10/1999) были приведены основные сведения, относящиеся к пространственной локализации единичного звукового источника в условиях отсутствия отражений (например, в безэховой камере или в большом хорошо заглушенном помещении). Эти данные уже достаточно широко известны в психоакустике и подробно освещены в литературе (например, в книге Блауэрта "Пространственный слух").

Однако для развития пространственных звуковых систем этой информации явно недостаточно. Поэтому в последние годы в разных странах проводятся многочисленные исследования по углубленному изучению возможностей слуховой системы в воссоздании пространственного звукового образа и оценке его тембральных характеристик. Результаты этих работ представлены в многочисленных докладах практически на всех последних конгрессах AES, на специальных конференциях, в статьях в таких журналах, как JAES, JASA, Acoustica и др.

Поскольку эти вопросы чрезвычайно важны для работы звукорежиссеров, инженеров, музыкантов и др, особенно в связи с освоением новых технологий записи и воспроизведения звука в системах Surround Sound, то представляется полезным рассказать о некоторых новых результатах, полученных в этом направлении за последнее время. Расскажу в основном об исследованиях, проводимых под руководством всемирно известных психоакустиков Б. Мура в Кембридже (Англия), Б Хартмана (США), а также о некоторых других исследованиях.

Большое практическое значение для развития пространственных аудиосистем имеют ответы на следующие вопросы:

- как слуховая система локализует звук в условиях, когда имеется много звуковых источников, работающих одновременно;
- как отражения в помещении влияют на пространственную локализацию;
- как в слуховой системе реализуется процесс локализации вообще.

Как уже было показано в вышеупомянутой статье, при локализации единичного источника точность локализации в горизонтальной плоскости достаточно высока и составляет примерно  $3^\circ$ , хотя имеются данные, что минимальное разрешение смещения источника может составлять даже  $1^\circ$ . Существуют два механизма локализации источника в горизонтальной плоскости:

- на низких частотах (до 1500 Гц) – это оценка разности по времени прихода звука к разным ушам (ITD). Интересно, что при угловом смещении в  $1^\circ$  разница по времени составляет  $\sim 10$  мс, что показывает очень высокую точность оценки в слуховой системе (рисунок 1);

- на высоких частотах (выше 2 кГц) – это оценка разности по интенсивности IID, возникающая за счет дифракции звука вокруг головы (рисунок 2).

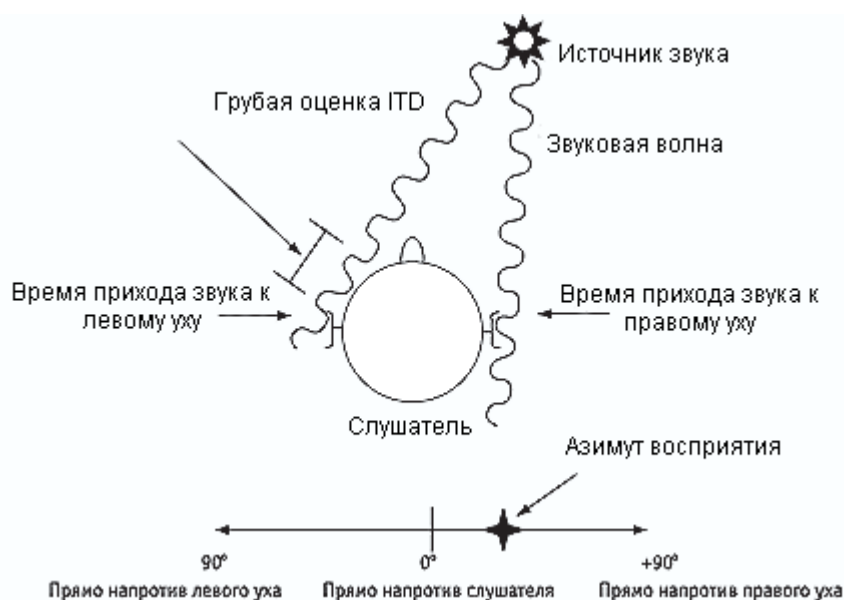


Рис.1 Локализация за счет разности во времени прихода звука-ITD

Локализация в вертикальной плоскости происходит за счет модификации спектра источника звука при взаимодействии с ушной раковиной, головой и торсом, которые действуют как частотно-зависимые дифракционные фильтры.

Анализ работы этих фильтров все время продолжается и уточняется, поскольку это очень важно для развития пространственных систем звуковоспроизведения и бинауральной стереофонии. Изменение формы АЧХ, измеренной в слуховом проходе, при перемещении единичного источника звука с равномерной АЧХ в вертикальной плоскости за счет дифракции на голове и ушной раковине показано на рисунке 3.

В целом точность локализации в вертикальной плоскости гораздо хуже, чем в горизонтальной, и составляет  $19-20^\circ$ , хотя минимально определяемый сдвиг составляет  $4^\circ$ , т.е. всего в четыре раза хуже, чем в горизонтальной плоскости.

Для ответа на первый поставленный выше вопрос – "Как происходит локализация, когда одновременно звучат несколько источников?", – были проведены многочисленные эксперименты, которые дали неутешительный ответ: точность локализации становится

значительно хуже, даже если работают только два источника с частично перекрывающимися

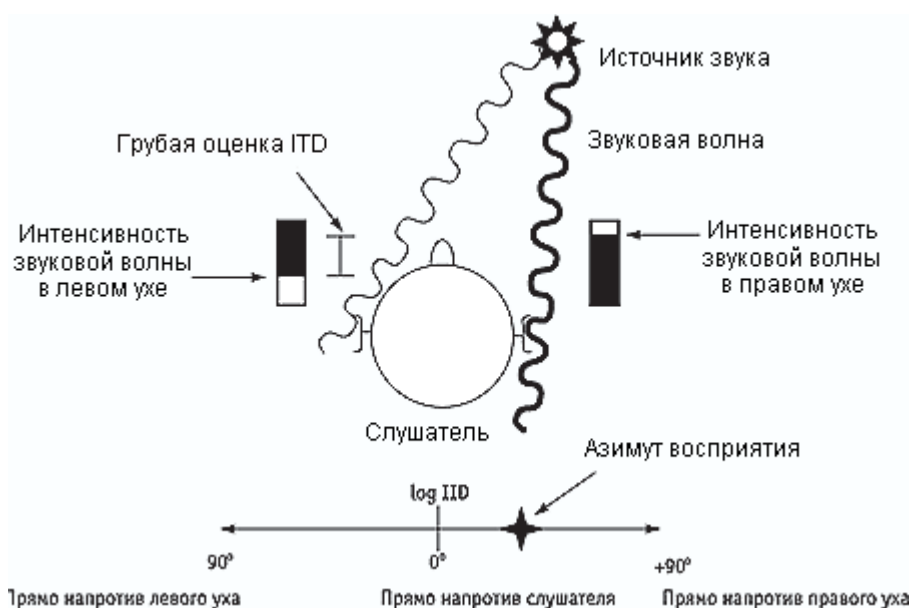


Рис.2 Локализация за счет разности интенсивностей-IID

спектрами – именно это и имеет место в стереосистемах, пространственных системах воспроизведения и др.

Например, были поставлены эксперименты со слушателями по локализации в заглушенной камере, где по периметру размещались источники звука (рисунок 4). Более "гуманные" опыты были выполнены с помощью "искусственной головы", которая размещалась в заглушенной камере, где по периметру в горизонтальной плоскости на расстоянии 4 м были установлены громкоговорители. На "искусственной голове" записывались импульсные характеристики в слуховом проходе. Затем звуки, предварительно обработанные в соответствии с записанными на "искусственной" голове импульсными характеристиками, подавались на стереотелефоны, и слушателей просили локализовать два одновременно звучащих источника, т.е. определить их угловое расстояние. Измерения были выполнены для разных сигналов: пары синусоид, различных видов модулированных сигналов и т.д. Если спектры сигналов, подаваемых из разных источников, перекрывались мало, то точность локализации составляла  $18^\circ$ , если спектры перекрывались, то  $60^\circ$ . Проще говоря, чтобы можно было определить, что звуки идут от разных источников, они должны были быть разнесены на  $60^\circ$ .

Это намного хуже, чем локализация одиночного источника. Кажется, что эти результаты противоречат практике – мы постоянно слышим несколько источников одновременно и обычно достаточно хорошо их локализуем. Однако реальные звуковые сигналы (музыка, речь, шум) обладают некоторыми особенностями, которые помогают слуховой системе.

Трудности, которые возникают при локализации нескольких источников, связаны, во-первых, с тем, что звуки от нескольких источников складываются в обоих ушах, и становится трудно определить разницу по времени и по интенсивности для каждого источника отдельно, чтобы установить локализацию каждого из них. Это можно сделать, если спектры звуков от каждого источника существенно отличаются (например, находятся в разных частях диапазона) – тогда слуховая система выполняет спектральный анализ, и в этом случае она справляется с задачей локализации.

Во-вторых, проблема состоит том, что перед слуховой системой встают две разные и конкурирующие между собой задачи: и локализовать звуки, и определить, к какому источнику они принадлежат (т.е. выполнить их сегрегацию, разделение на звуковые потоки).

В моей статье в 4/2001 уже было немного сказано о механизме сегрегации, сейчас поговорим об этом подробнее.

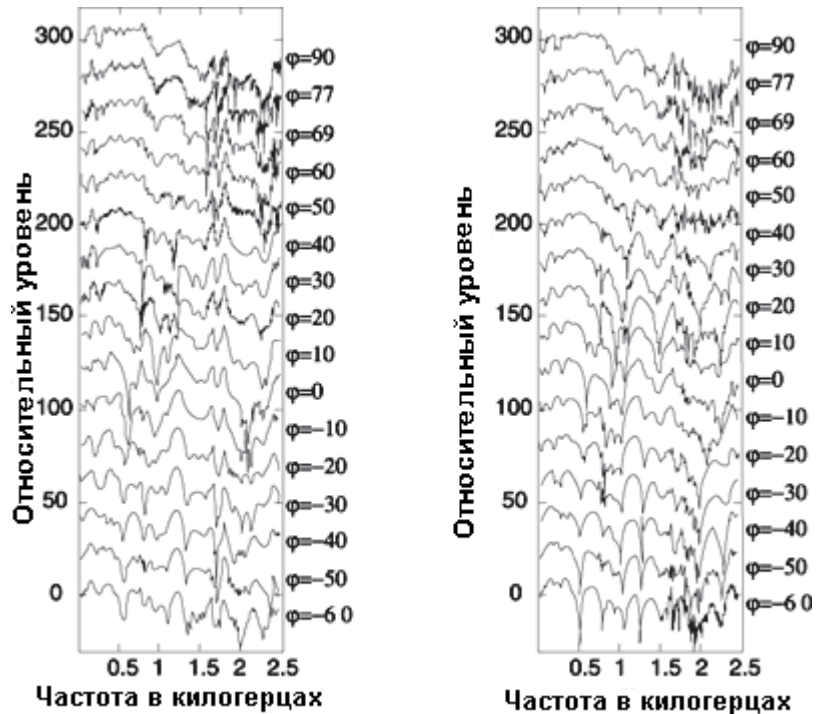


Рис.3 Форма АЧХ, измеренная в левом и правом слуховых каналах при разных углах подъема звукового источника

Одними из очень важных критериев объединения звуков в один звуковой поток, то есть приписывания их одному источнику, являются подобие спектров и характер переходных процессов, а также их синхронизация по времени – если звуки включаются и выключаются одновременно, то слуховая система обычно считает, что они исходят от одного источника, даже если на самом деле они разнесены в пространстве. Это очень важный вывод для пространственной звукозаписи. Реальные источники редко включаются и выключаются синхронно, кроме того их спектры постоянно меняются во времени – какие-то источники (инструменты) доминируют в разные моменты времени, поэтому слух успевает их выделить и локализовать.

В связи с этим возникают чрезвычайно интересные вопросы: "Зависит ли процесс сегрегации от процессов локализации звуков?" "Какой из процессов происходит сначала?" "Используются ли в процессе сегрегации какие-то признаки, по которым слуховая система локализует звуки?"

Обычно в процессе прослушивания звуков от различных источников (например, инструментов в оркестре) не возникает особых трудностей выделения отдельных источников звука. Насколько хорошо это делается, зависит от состояния слуха и степени музыкальной подготовленности слушателя. Когда в слуховую систему поступают звуки от двух звуковых источников, например от скрипки и рояля, то в высших отделах коры головного мозга (а не в периферической слуховой системе) происходит разделение всех поступивших звуков на два отдельных потока – один относится к роялю, другой – к скрипке. При этом происходят два разных процесса: один – разделение одновременно поступающих звуков, другой – последовательное разделение их во времени. Эти два процесса называются последовательной и параллельной группировкой (сегрегацией).

Для такого разделения используется большое количество различных физических признаков (часть из них была упомянута выше). Эксперименты показали, что некоторые из признаков, по которым система производит локализацию в пространстве, (например разница во времени прихода звуков к разным ушам – ITD), оказывают существенное влияние и на процесс последовательной группировки – но оказывают относительно малое влияние на процесс одновременной группировки. Это подтверждается различными экспериментами: например, через головные телефоны одновременно предъявлялись разные гласные звуки, и слушателей просили их различить. Если между звуками вводилась задержка во времени 400 мс (что соответствует углу между ними в  $45^\circ$ , как если бы они воспроизводились через разные громкоговорители), то различимость улучшалась всего на 7%, в то же время разница в полтона по основной частоте улучшала их различимость на 22%. Однако при предъявлении последовательных звуков введение дополнительной временной задержки (ITD) существенно улучшало их различимость.

Результаты исследований позволяют сделать важный для практики вывод: при прослушивании одновременных звуков от разных источников слуховая система сначала производит их группировку по потокам (определяет, какие звуки к какому источнику принадлежат), используя при этом различные непространственные признаки (значение основной частоты, степень гармоничности, амплитудную огибающую, структуру переходных процессов и др), а затем локализует данный источник звука, т. е. сначала определяет, "Что это", а затем – "Где это" (рисунок 5).



*Рис.4 Размещение громкоговорителей на сфере в заглушенной камере*

Затем слух продолжает разделять последовательно поступающие звуки по потокам (каждому инструменту – свой), но при этом уже использует пространственные признаки: разницу во времени поступления, разницу в амплитудах в разных ушах и др. Интересно, что если при этом источник звука не меняет своего положения или изменяет его плавно (например, солист на сцене), то слуховая система продолжает воспринимать этот звук как часть одного слухового потока, то есть считает исходящим от единого звукового источника. Но если источник звука резко меняет свое положение в пространстве (локализацию), то слуховая система может воспринять его как совершенно другой источник звука – процесс сегрегации произведет расщепление звукового потока.

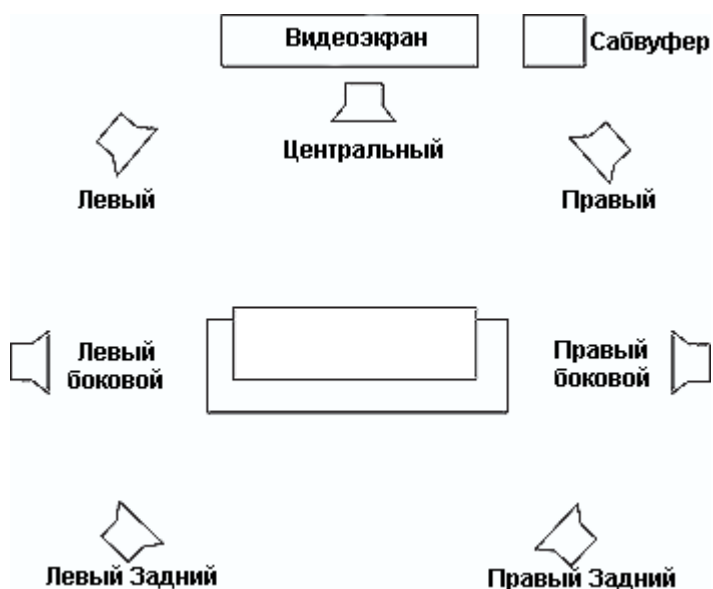


Рис.5 Процесс сегрегации и локализации

Учитывая сложность одновременно действующих процессов при восприятии нескольких звуковых источников, разнесенных в пространстве, например, в системах Surround Sound, неудивительно, что точность локализации при этом ухудшается.

Следующая проблема, которая возникает при локализации нескольких источников звука, состоит в оценке влияния на точность локализации отраженных сигналов, которые возникают в любом помещении прослушивания (студии, концертном зале, комнате прослушивания). Как известно (см. "Звукорежиссер" 10/2000), структура отраженных сигналов в помещении имеет вид, представленный на рисунке 6. Если источник звука излучает короткий импульс, то к слушателю (или микрофону), находящемуся в определенном месте помещения, сначала приходит прямой сигнал, затем, через определенные отрезки времени, первые отраженные сигналы (как правило, от потолка, боковых стен, пола), затем количество этих отраженных сигналов возрастает, и процесс спада уровня сигнала становится почти непрерывным. Время, в течение которого сигнал спадает на 60 дБ, называется "временем стандартной реверберации".

Процесс реверберации определяет качество тембра воспринимаемого звука в помещении, кроме того, он оказывает существенное влияние на процесс локализации источников звука в помещении. Наличие отраженных звуков, приходящих со всех направлений, создает случайные вариации признаков, определяющих локализацию, и неизбежно ухудшают ее точность. Особенно страдает от этого такой признак локализации, как разница во времени между сигналами на двух ушах (ITD). В помещении, где отраженные звуки преобладают над прямыми, этот критерий локализации вообще становится ненадежным. Зато второй критерий, разница по интенсивности (IID), "страдает" меньше, так как он используется на высоких частотах, где коэффициент поглощения на поверхности помещения значительно возрастает с частотой, и уровень отраженных сигналов уменьшается. Если локализуются широкополосные сигналы от разных источников в сильно реверберирующем помещении, то слух, в основном, полагается на информацию от высокочастотной части спектра, используя только второй критерий (IID). Такая переоценка локализационных признаков происходит на подсознательном уровне.



Рис.6 Структура отраженных сигналов в помещении

Второй механизм, помогающий осуществлять локализацию источников звука в помещении, – "эффект предшествования", или "эффект Хааса", или "закон первой волны". Явление это известно достаточно давно, однако его объяснение с точки зрения современной психоакустики появилось только в настоящее время.

Сущность этого явления заключается в следующем: если звуки с коротким интервалом задержки по времени приходят с разных направлений, то локализация общего звука происходит по первому пришедшему звуку, т.е. слуховая система как бы теряет способность локализовать отраженный звук, если он приходит через слишком короткий отрезок времени (рисунок 7).

Слуховая система отдает предпочтение первому "прямому" звуку, который несет более точные данные о локализации источника по сравнению с отраженным звуком, который искажает информацию о локализации. Это своего рода "нейронные ворота", которые открываются в момент атаки звука, производят его локализацию и закрываются.

Нужно сказать, что этот отраженный звук все-таки оказывает свое влияние на точность локализации первого основного звука – если место появления отраженного звука все больше отодвигается от расположения прямого звука, то он как бы "утягивает" локализацию первого звука за собой примерно на  $7^\circ$  (меняется протяженность первого источника); при большем смещении эффект уже не сказывается.

Если интервал между двумя короткими звуками становится слишком коротким (меньше 1 мс), то эффект предшествования не проявляется, происходит некоторая компромиссная (усредненная) локализация. Этот эффект называется "суммарная локализация". Если интервал больше 5 мс для импульсов (щелчков) и больше 40 мс для речи и музыки, то слышны отдельно и прямой звук, и эхо, то есть эффект предшествования также не проявляется.

Если уровень отраженного звука сделать на 10...15 дБ больше уровня прямого звука, то эффект предшествования также пропадает – слышны два разных звука с разных направлений.

Этот эффект проявляется обычно, когда два звука похожи по спектру, что и имеет место в прямом и отраженном сигнале.

Однако эксперименты показали, что эффект имеет место и для двух разных звуков, например, прямой низкочастотный сигнал очень эффективно подавляет локализацию высокочастотного отраженного сигнала.

Интересно отметить, что этот эффект проявляется не только тогда, когда сигналы приходят из разных направлений в горизонтальной плоскости, где основную роль играет разница по времени и интенсивности. Эффект Хааса имеет место и при локализации прямого и отраженного звуков в вертикальной плоскости, правда, он выражен значительно слабее.

Необходимо сказать, что эффект предшествования не подавляет всю информацию об отраженном сигнале – слушатель легко различает разницу в тембрах прямого звука и звука, дополненного отражениями. Эта тембральная разница несет информацию о размерах помещения, позиции стен, потолка и др. Эффект предшествования проявляется только в том, что ранние отраженные сигналы не слышны как отдельные звуки, и информация об их пространственной локализации теряется.

Таким образом, точность локализации источников звука в помещении при наличии отражений существенно ухудшается по обычным критериям (ITD, IID). Однако слух использует два других механизма – локализацию по IID в высокочастотной части диапазона и эффект предшествования, что позволяет осуществлять локализацию, хотя и с меньшей точностью.

Одни из самых последних исследований в психоакустике посвящены третьей проблеме: "Как в слуховой системе реализуется процесс подавления первых отражений и процесс локализации вообще?". Является ли это следствием работы специализированных нейронов, или это продукт сознательного принятия решений высшими отделами головного мозга?

Исследования на животных позволили выявить специальные бинауральные нейроны, способные сравнивать сигналы от двух ушей и реагировать на разницу во времени и на разницу в интенсивности между ними. На нейронах в этих же отделах мозга было выявлено, что при подаче двух коротких щелчков с изменяемой задержкой между ними чувствительность нейронов ко второму звуку при коротких задержках подавляется. Что касается реакции

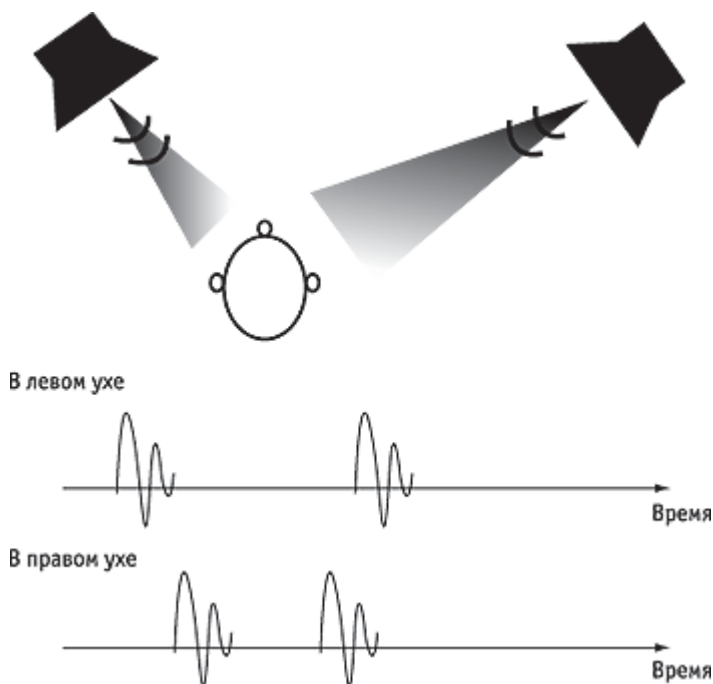


Рис.7 Эффект предшествования

человека, то, если бы эффект предшествования определялся только реакцией нейронов, он происходил бы практически мгновенно. Однако выяснилось, что он требует определенного времени для возникновения, то есть слуховая система как бы "обучается".

Например, были проделаны такие эксперименты: если подать два сигнала с задержкой 8 мс (что моделирует как бы прямой звук и его задержанное эхо), то в первый момент эти два сигнала слышны раздельно, но если их повторить несколько раз, например со скоростью четыре раза в секунду, то через некоторое время второй звук перестает быть слышимым. Эффект предшествования может быть разрушен резким изменением акустической обстановки: если один сигнал подавать от одного громкоговорителя, а другой с некоторой задержкой от другого, то после определенного периода обучения возникает эффект подавления, но, если внезапно изменить расположение громкоговорителей (или одного из них), то эффект пропадает, и каждый звук слышен отдельно.

Все эти эксперименты заставляют предположить, что восприятие эффекта предшествования является актом сознания, а не физиологической особенностью. Создается впечатление, что эффект предшествования срабатывает только тогда, когда время появления эха, его амплитуда и направление совпадают с некоторым "ожиданием" слушателя от акустики данного помещения. Это ожидание формируется на основании предшествующего опыта прослушивания в данном зале (или подобных), зрительного впечатления, предварительного обучения и др. Однако быстрое изменение позиции ведущего и ведомого звука, несовпадающего со слушательскими ожиданиями от акустики данного помещения, сразу же делают эхо слышимым, то есть нарушают эффект. Аналогичные результаты получаются при изменении спектра эха или направления его прихода, которые, по мнению слушателя, делают его неестественным для данного помещения, что также делает его слышимым.

Таким образом, как только нарушаются траектории прихода звука и его параметры, выстроенные слушателем в сознании при предварительном анализе акустики данного помещения, так эффект предшествования сразу пропадает.

Создается впечатление, что сначала прямой звук и эхо-сигнал обрабатываются слуховой системой совместно, обогащая спектр (тембр) прямого звука, затем оба сигнала обрабатываются высшей нервной системой, и она принимает решение – подходит ли данный звук по своим параметрам на роль эха от прямого сигнала в данном помещении. Если подходит, то информация о нем подавляется, и локализация происходит только по прямому звуку; если не подходит, то он слышен как отдельный звук, и происходит локализация двух разных источников.

Следовательно, при локализации звуков от разных, одновременно работающих источников в реальном помещении (например, в системах домашнего театра) слуховой системе в процессе локализации, то есть в процессе построения пространственного образа, приходится решать сразу две различные и трудные задачи:

- произвести классификацию всех поступивших звуков по потокам, определить к какому источнику какой звук принадлежит;
- произвести локализацию всех источников, для чего необходимо еще решить проблему, какие звуки идут прямо от источников, а какие являются их отраженными сигналами (эхо), и поэтому их можно не принимать во внимание при локализации.

Неудивительно, что при таких условиях общая точность локализации снижается, и, соответственно, увеличивается время ее реализации.



Восприятие пространственных звуковых систем зависит не только от точности локализации, но и от особенностей восприятия тембра пространственного слухового образа. Об этом речь пойдет в следующей части этой статьи.