

ОПИСАНИЕ ОБЛАСТИ ПРИМЕНЕНИЯ

Измерение разборчивости речи в ПО DIRAC, тип 7841

Разборчивость речи – это важная характеристика эффективности и качества систем связи или передачи в зашумленной обстановке. В повседневной жизни требуется понять, что было сказано, например, через громкоговоритель, а также правильно среагировать на различные сигналы.

Системы усиления речи, такие как системы оповещения, телефоны, радио, переговорные устройства очень важны для общества, однако, в офисах, на рабочем месте, в автомобилях и во многих других случаях необходимо создать условия для простой, не усиленной речи.

Современное оборудование компании Brüel & Kjær позволяет выполнять измерение речи при помощи направленного громкоговорителя типа искусственный рот или путем прямой инъекции в звуковую систему с учетом воздействия внешних и внутренних источников шума. Объективная оценка может быть выполнена при помощи одного или нескольких стандартизованных методов, которые позволяют определить причину низкой разборчивости речи.



ПО DIRAC

DIRAC – это программное обеспечение для ПК, предназначенное для измерения широкого набора акустических параметров помещения. Измерение разборчивости речи может быть выполнено в соответствии с требованиями стандарта МЭК 60268-16 при помощи направленного громкоговорителя типа искусственный рот или путем прямой инъекции в звуковую систему, учитывая при этом воздействие фонового шума. Для точного измерения по стандарту ISO 3382 могут применяться внутренние или внешние, созданные при помощи громкоговорителя, возбуждающие сигналы.

Акустические решения компании Brüel & Kjær для помещений позволяют выполнить следующие измерения:

- определить достаточность разборчивости речи, например, в церкви или на железнодорожной станции;
- определить количество громкоговорителей, необходимое для увеличения разборчивости речи до приемлемого уровня;
- определить ситуации, в которых разборчивость речи определяется реверберацией и фоновым шумом;
- определить насколько хорошо актер на сцене слышит себя и других;
- оптимизировать акустические параметры помещения, чтобы улучшить разборчивость речи.

ПО DIRAC – это ценный инструмент не только для инженеров-акустиков, работающих в лабораториях и в полевых условиях, он может применяться также в исследовательских и образовательных институтах. Чтобы получить дополнительную информацию об области применения ПО DIRAC, необходимо посетить сайт [visit www.bksv.com/DIRAC](http://www.bksv.com/DIRAC).

Основы разборчивости речи

Речь, предназначенная для другого лица в помещении или передающаяся по системе оповещения пассажиров, никогда не принимается слушателем в том виде, в котором она была произнесена. Разборчивость речи в повседневной обстановке зависит от следующих факторов:

- уровень фонового шума;
- расстояние от говорящего до слушателя;
- громкость речи (мощность сигнала);
- спектр голоса речи;
- уровень реверберации (эхо).

На разборчивость речи негативно влияет шум. Чтобы добиться полной разборчивости предложения слушателем с нормальным слухом, отношение сигнал/шум – разница между уровнем звука и уровнем мешающего шума – должно быть не менее 15 дБ (А). Качество звука речи определяется механизмом снижения разборчивости. Гласные и согласные звуки обладают разной энергией. Согласные произносятся более мягко, чем гласные, поэтому они легче тонут в шуме. Средний уровень согласных на 10-12 дБ меньше уровня гласных.

Таблица 1.

Сила голоса
и уровень звука

Сила голоса	дБ (А)
Шепот	32
Мягкий	37
Расслабленный	42
Нормальный (личная беседа)	47
Нормальный (публичное выступление)	52
Повышенный	57
Громкий	62
Очень громкий	67
Крик	72
Максимально сильный крик	77

Для понимания речи согласные играют большую роль, чем гласные. Они передают больше словесной информации, выполняя функцию точек перехода и разделения между звуками и словами. Кроме того, большая часть энергии согласных находится в области высоких частот, в то время как гласным соответствуют средние и низкие частоты.

Звуковой сигнал речи богат характерными временными и частотными паттернами. Речь охватывает диапазон частот от 125 Гц до 8 кГц, а ее наиболее важная часть лежит в диапазоне от 300 до 3000 Гц. Амплитуда огибающей сигнала речи изменяется в широком диапазоне. Такое изменение мощности по времени и частоте называется модуляцией. Разборчивость речи в большой степени определяется этой медленной модуляцией сигнала звукового давления, невзирая на его акустическую сложность. Произнесенные слова остаются разборчивыми даже в случае сильной деградации сигнала как во временной, так и в частотной области.

Объективные методы и параметры

Разборчивость речи, как правило, выражается в процентном отношении слов, предложений или фонем (произносимые звуки, из которых состоит слово), правильно распознанных слушателем или группой слушателей и произнесенных оратором или группой ораторов. Однако такая субъективная проверка разборчивости занимает много времени и стоит достаточно дорого. К счастью, оценить разборчивость речи можно при помощи физических измерений, не требующих привлечения к испытаниям живых объектов. Такие объективные методы оценки разборчивости речи описаны стандартами Международной электротехнической комиссии (МЭК) и Международной организации по стандартизации (ISO).

В методе физического измерения применяется синтезированный сигнал, аналогичный речи и воспроизводимый таким устройством, как искусственный рот или небольшой громкоговоритель, который в испытании играет роль источника сигнала. Такой сигнал может быть также воспроизведен через систему громкой связи или систему создания звукового поля. Микрофон, находящийся в месте расположения слушателя, принимает этот сигнал, после чего вычисляется степень ухудшения модуляции в каждой полосе частот полученной речи. Шум и реверберация помещения снижают качество модуляции и сказываются на результатах.

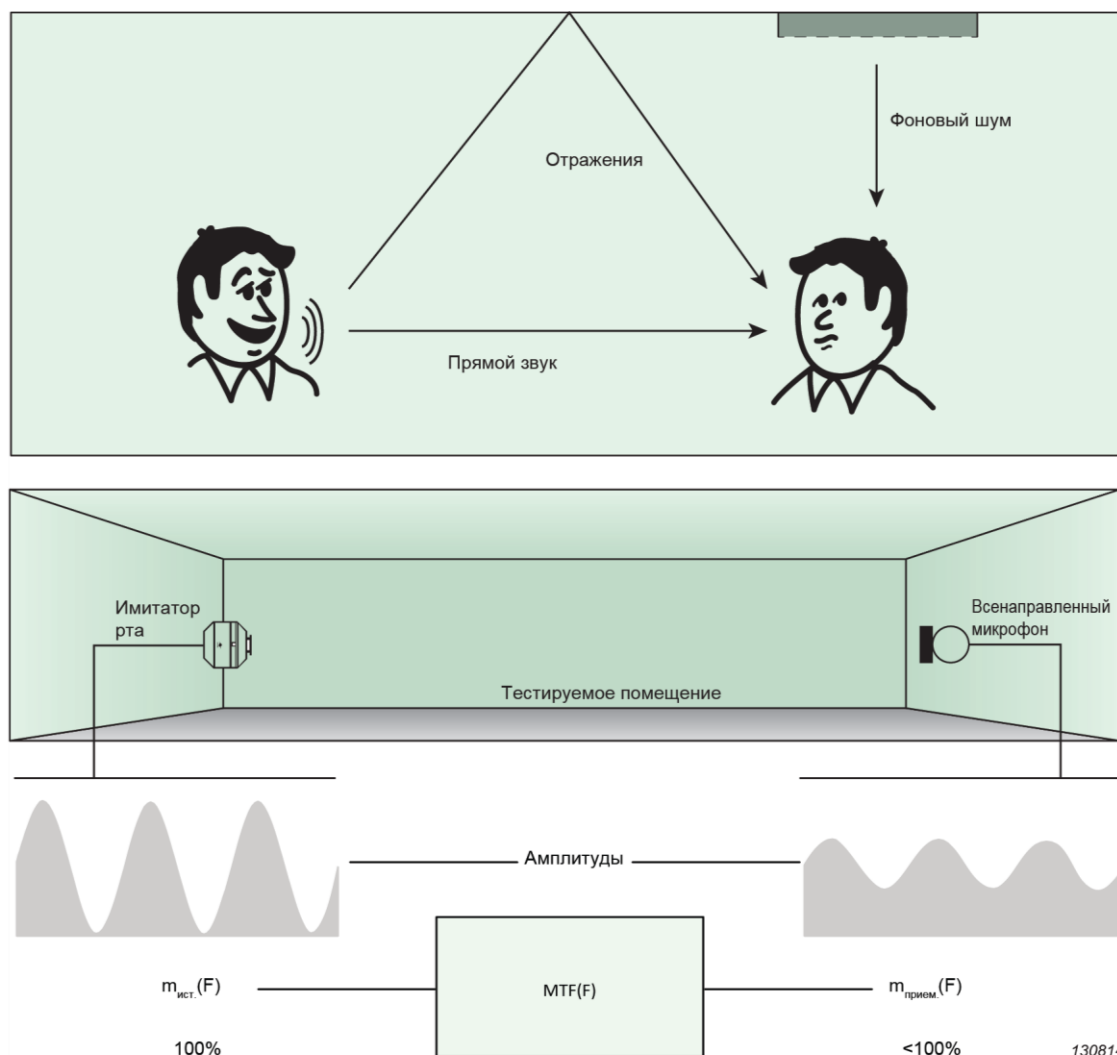
Для измерения разборчивости речи используются широко распространенные параметры: STI, RASTI или STIPA. Все эти параметры имеют один и тот же смысл, однако параметр RASTI и STIPA являются упрощенной версией параметра STI. Все они основаны на измерении функции MTF (модуляционно-передаточная функция – от англ. Modulation Transfer Functions) в семи октавных полосах частот. Для каждого октавного диапазона вычисляется одна функция MTF, которая позволяет численно оценить степень сохранения амплитуды модуляций в этом диапазоне. Эти функции показывают, как сильно сохраняется амплитудная модуляция в семи октавных диапазонах спектра продолжительной речи.

Типовое измерение разборчивости речи может быть выполнено следующим образом:

1. сигнал возбуждения воспроизводится при помощи направленного источника звука типа искусственный рот, который устанавливается в месте расположения оратора. Сигнал с экспоненциальным изменением частоты (E-sweeps) является более неприятным для слуха по сравнению с последовательностью импульсов максимальной длины (MLS), звук которых менее навязчив;
2. отклик регистрируется либо непосредственно в ПО DIRAC с целью немедленного получения результатов, либо записывается при помощи портативного регистратора звука, который более удобен в эксплуатации;
3. по возвращении в офис можно посмотреть изменение разборчивости речи в случае изменения уровня сигнала и/или уровня фонового шума.

Параметры разборчивости речи в ПО DIRAC основаны на взаимосвязи между воспринимаемой разборчивостью речи и амплитудной модуляцией в голосе говорящего, как описал Хогаст (Houtgast) и др. [1, 6]. Если источник звука в помещении создает шум, который обладает низкочастотной синусоидальной амплитудной модуляцией на низкой частоте, глубина модуляции в точке приема будет снижена из-за наличия в помещении отражений и фонового шума. Модуляционно-передаточная функция (MTF) описывает, в какой степени модуляция передается от источника приемнику, как функция от частоты модуляции F , которая лежит в диапазоне от 0,63 до 12,5 Гц. При этом функция MTF зависит от свойств системы и фонового шума, см. рис. 1.

Рисунок 1.
Зависимость между разборчивостью речи и глубиной модуляции



Матрица снижения модуляции

Чтобы оценить разборчивость речи, функция MTF определяется для каждого частотного диапазона речи. Высокие значения функции MTF говорят о хорошей передаче уровня модуляции и соответствуют хорошей разборчивости речи. Низкие значения функции MTF говорят о значительном снижении разборчивости речи, вызванном акустическими свойствами системы и/или фоновым шумом.

Результат усреднения значений MTF на 14 частотах модуляции называется индексом передачи модуляции (MTI – от англ. modulation transmission index). Индексы передачи модуляции для семи октавных диапазонов могут быть обработаны с целью получения индекса передачи речи (STI – от англ. Speech Transmission Index), см. также [1]. В табл. 2 показан пример матрицы снижения модуляции.

По данным этой матрицы можно получить информацию о причине снижения разборчивости речи. Постоянное значение MTF с увеличением частоты F говорит о наличии фонового шума, постоянное возрастание величины MTF указывает на наличие реверберации, а уменьшение и последующее увеличение величины MTF является признаком эха. Например, матрица снижения модуляции, представленная в табл. 2, показывает случай, когда подавляющее влияние оказывает реверберация, а фоновым шумом и эхом можно пренебречь.

Таблица 2.
Пример матрицы снижения модуляции STI

Частота модуляции F, Гц	Октавный диапазон частот, Гц						
	125	250	500	1 к	2 к	4 к	8 к
0,63	1,000	1,000	1,000	1,000	1,000	1,000	1,000
0,80	1,000	1,000	1,000	1,000	1,000	1,000	1,000
1,00	1,000	1,000	1,000	1,000	1,000	1,000	1,000
1,25	1,000	1,000	1,000	1,000	1,000	1,000	1,000
1,60	0,858	0,866	0,806	0,852	0,837	0,841	0,835
2,00	0,858	0,866	0,806	0,852	0,837	0,841	0,835
2,50	0,858	0,866	0,806	0,852	0,837	0,841	0,835
3,15	0,858	0,870	0,806	0,852	0,837	0,841	0,835
4,00	0,858	0,866	0,806	0,852	0,837	0,841	0,835
5,00	0,651	0,676	0,543	0,664	0,630	0,633	0,612
6,30	0,651	0,676	0,543	0,664	0,630	0,633	0,612
8,00	0,506	0,533	0,377	0,531	0,506	0,502	0,471
10,00	0,506	0,533	0,377	0,531	0,506	0,502	0,471
12,50	0,444	0,460	0,267	0,427	0,406	0,413	0,373
MTI	0,75	0,76	0,70	0,75	0,73	0,74	0,73

Используемые октавные частотные диапазоны соотносятся с типовым частотным диапазоном человеческого голоса. Разница между спектром мужского и женского голоса представлена в стандарте МЭК 60268-16 [2]. Спектр женского голоса не имеет октавного диапазона 125 Гц.

Измерение функции MTF: модулированный шум или импульсная характеристика

Для измерения функции MTF широко применяется два метода: метод модулированного шума и метод импульсной характеристики.

В методе модулированного шума сигнал возбуждения состоит из $7 \times 14 = 98$ суммируемых сигналов шума, каждый из которых фильтруется и модулируется в соответствии с матрицей, представленной в табл. 2. Сигнал измеряется в точке расположения слушателя, а функция снижения модуляции MTF (F) измеряется для каждого октавного диапазона частоты модуляции (F). Трудность полного измерения индекса STI по методу, описанному здесь, заключается в том, что модуляция в одном октавном частотном диапазоне может влиять на модуляцию в других частотных диапазонах. Следовательно, не все 98 могут быть измерены одновременно. Кроме того, неслучайный характер сигнала возбуждения требует большого времени для получения воспроизводимых результатов. На практике для одного полного измерения индекса STI требуется не менее 15 минут. Флуктуация фонового шума может восприниматься в приемнике как флуктуация модулированного сигнала, что может увеличить разборчивость речи при низком отношении сигнал/шум.

Schroeder [3] показал, что функция MTF может быть получена при помощи преобразования Фурье от отклика на воздействие прямоугольным импульсом. Райф (Rife) [4] при помощи [1] и [3] учитывает вклад фонового шума. Если импульсная характеристика измеряется путем обратной свертки детерминированного сигнала, такого как последовательность импульсов максимальной длины (MLS) или сигнал с изменением частоты, измерение при той же повторяемости занимает меньше времени, чем применение по методу модулированного шума (около 5 с на усреднение). Однако метод импульсной характеристики требует больше вычислительной мощности.

Условия и ограничения для измерения функции MTF

Функция MTF, являющаяся базисом для определения разборчивости речи, также имеет ограничения. Искажения, возникающие в тестируемой системе, могут влиять на функцию MTF (и, следовательно, на измерение разборчивости речи), что делает измерение отличным от реальной разборчивости. Например, записанный голос, воспроизводимый со слегка завышенной скоростью, становится очень разборчивым, однако при этом функция MTF значительно снижается. Ограничение сигнала по амплитуде (перегрузка) может сказаться на разборчивости речи значительно сильнее, чем это покажет функция MTF. Аналогично ведет себя сигнал с провалами (Хогаст (Houtgast), Стинекен (Steeneken) и др. [6]). В общем случае, различие между реальной и измеренной разборчивостью речи для двух упомянутых ранее методов измерения будет различным.

В стандарте МЭК 60268-16 [2] рассматриваются условия, позволяющие избежать этой проблемы:

1. тестируемая система не должна создавать смещение частоты или использовать частотное умножение;
2. тестируемая система не должна содержать вокодеров, таких как LPC, CELP и RELP;
3. передача речи должна иметь хорошую линейность, амплитудная компрессия или расширение должны быть ограничены величиной 1 дБ, а сигнал не должен быть ограничен по амплитуде.

Таким образом, можно заключить:

4. тестируемая система не должна приводить к ограничению сигнала по амплитуде;
5. системе не должна приводить к образованию провалов.

Другими словами, во время измерения разборчивости речи в звуковой системе важно избегать любого нелинейного поведения. Если система ведет себя линейно, измеренная и реальная разборчивость речи очень хорошо коррелируют между собой для обоих методов.

Параметры, связанные с разборчивостью речи

Индекс передачи речи (STI)

Индекс передачи речи (STI) является наиболее исчерпывающим и важным параметром разборчивости речи в ПО DIRAC. Несмотря на то, что он неприменим для каналов передачи, в которых имеется смещение или умножение частоты, а также присутствуют вокодеры (речевые кодеры), индекс STI учитывает большое количество факторов, которые определяют разборчивость речи. Более подробно см. в [2].

Технически, индекс STI вычисляется как взвешенная сумма индексов передачи модуляции MTI по каждому октавному диапазону от 125 Гц до 8 кГц (где каждое значение MTI определяется из функции MTF на 14 различных частотах модуляции, см. табл. 2), учитывая влияние аудитории в соответствии со стандартом МЭК 60268-16.

Таблица 3.
Зависимость
между индексом
STI и разборчи-
востью речи

STI	0,00 – 0,30	0,30 – 0,45	0,45 – 0,60	0,60 – 0,75	0,75 – 1,00
Разборчивость речи	Ужасная	Плохая	Удовлетворительная	Хорошая	Замечательная

Параметр STI, описанный в этом разделе, основан на определениях, данных в стандарте МЭК ЕС 60268-16 ред. 4.0. Стандарт ISO 3382-3 для офисов открытого плана [7] использует несколько другие определения индекса STI, в которых не учитывается маскирование аудитории и порог слышимости. Кроме того, индекс STI стандарта ISO 3382-3 основан на другом (независимым от пола) спектре речи.

ПО DIRAC имеет отдельный параметр “STI for ISO 3382-3” (Индекс STI для стандарта ISO 3382-3).

Индекс передачи речи с учетом акустики помещения (RASTI)

Индекс передачи речи с учетом акустики помещения RASTI (от англ. Room Acoustics Speech Transmission Index) предназначен для адаптации индекса STI под типовые акустические условия в помещении. Индекс RASTI разработан специально с целью уменьшения времени измерения (методом модулированного шума) и времени получения конечного результата. Чтобы вычислить индекс RASTI правильно, кроме требований, предъявляемых к методу STI, должно выполняться условие равномерности частотной характеристики всей системы в октавных диапазонах от 125 Гц до 8 кГц, фоновый шум должен быть гладким как по времени, так и по частоте, пространство не должно иметь дискретных источников эха, а время реверберации не должно жестко зависеть от частоты. Более подробно см. в [2].

Индекс RASTI вычисляется как взвешенная сумма индексов MTI в октавных диапазонах с 500 до 2000 Гц, где значения MTI вычисляются из функции MTF на 4 и 5 различных частотах модуляции соответственно. См. табл. 4.

Таблица 4.
Матрица сниже-
ния модуляции
RASTI

Частота модуляции F, Гц	Октавный диапазон частот, Гц	
	500	2 к
0,7		•
1,0	•	
1,4		•
2,0	•	
2,8		•
4,0	•	
5,6		•
8,0	•	
11,2		•

Для получения индекса STI в ПО DIRAC отдается предпочтение методу импульсной характеристики, а не методу модулированного шума, поэтому индекс RASTI не имеет преимуществ по сравнению с индексом STI, если рассматривать время измерения и время вычисления. Несмотря на это, индекс RASTI требует наличия измерительной системы с меньшей полосой частот по сравнению с индексом STI и может применяться для обзорного измерения акустики в большинстве встречающихся на практике помещений. Параметр RASTI признан устаревшим в стандарте МЭК 60268-16 ред. 4.0. Поэтому им не следует больше пользоваться.

Индекс передачи речи для систем телекоммуникации (STITEL)

Индекс передачи речи для систем телекоммуникации STITEL (от англ. Speech Transmission Index for Telecommunication Systems) также требует меньшего времени измерения и вычисления. Чтобы индекс STITEL соответствовал индексу STI, необходимо, чтобы выполнялся ряд условий измерения, которые являются типовыми для систем телекоммуникации.

Для вычисления индекса STITEL используются те же октавные диапазоны, что и для индекса STI, однако в каждом из них используется только одна определенная частота модуляции. См. табл. 5.

Таблица 5.
Снижение модуляции STITEL

Частота модуляции F, Гц	Октавный диапазон частот, Гц						
	125	250	500	1 к	2 к	4 к	8 к
0,71			•				
1,12	•						
1,78						•	
2,83				•			
4,53							•
6,97					•		
11,33		•					

Для получения индекса STI в ПО DIRAC отдается предпочтение методу импульсной характеристики, а не методу модулированного шума, поэтому индекс STITEL не имеет преимуществ по сравнению с индексом STI, если рассматривать время измерения и время вычисления, но его можно использовать для сравнения с другими измеренными значениями STITEL.

Индекс передачи речи для систем оповещения (STIPA)

Индекс передачи речи для систем оповещения STIPA (от англ. Speech Transmission Index for PA Systems) является упрощенной версией индекса STI и предназначен для его имитации в условиях, типовых для систем оповещения. Индекс STIPA разработан с целью уменьшения времени измерения (методом модулированного шума) и времени получения конечного результата. Чтобы индекс STIPA соответствовал индексу STI, необходимо, чтобы выполнялся ряд условий измерения, которые являются типовыми для систем оповещения.

Таблица 6.
Матрица снижения модуляции STIPA

Частота модуляции F, Гц	Октавный диапазон частот, Гц						
	125	250	500	1 к	2 к	4 к	8 к
0,63			•				
0,80						•	
1,00		•					
1,25					•		
1,60	•						
2,00				•			
2,50							•
3,15			•				
4,00						•	
5,00		•					
6,25					•		
8,00	•						
10,00				•			
12,50							•

Для получения индекса STI в ПО DIRAC отдается предпочтение методу импульсной характеристики, а не методу модулированного шума, поэтому индекс STIPA не имеет преимуществ по сравнению с индексом STI, если рассматривать время измерения и время вычисления, но его можно использовать для сравнения с другими измеренными значениями STIPA.

Потеря согласных из-за артикуляции, выраженная в процентах (% ALC)

Выраженный в процентах параметр потери согласных из-за артикуляции % ALC (от англ. Percentage Articulation Loss of Consonants), обозначаемый также как % ALcons, основан на восприятии слов слушателем. В ПО DIRAC параметр % ALC вычисляется из индекса STI по широко распространенной формуле приближения Фарелла-Беккера (Farrel Becker)^{*1} [8]:

$$\% \text{ ALC} = 170,5405 e^{-5,419(\text{STI})}$$

^{*1} Редактор [6], см. прим. на стр. 81: это выражение, известное как формула Фарелла-Беккера, часто используется для перевода значения STI в значение ALcons. Похоже, что источник этого уравнения не публиковался в открытой литературе. Тем не менее, между этой формулой и эмпирическими данными, полученными Хогастом и др. (по рисунку, а не по уравнению), наблюдается соответствие [1]. Можно заключить, что это выражение было получено либо путем аналогичных экспериментов, либо было выведено из данных, полученных Хогастом и др.

Аналогичная формула используется для RASTI, STITEL и STIPA. Значения % ALC, как правило, лежат от 0 (соответствует исключительно хорошей разборчивости речи) до 100 (чрезвычайно плохая разборчивость), однако значение % ALC, вычисленное по упомянутой выше приближенной формуле, может превышать величину 100 при очень низком индексе STI.

В ПО DIRAC значение % ALC используется, главным образом, для сравнения с другими вычисленными или измеренными значениями % ALC.

Отношение сигнал/шум

Отношение сигнал/шум SNR (от англ. Signal-to-Noise Ratio) определяется как логарифмическое отношение уровня сигнала к уровню шума и поэтому имеет большее отношение к сигналу, а не к системе. В ПО DIRAC отношение сигнал/шум (SNR) определяется как отношение речи к фоновому шуму, вычисленное при помощи импульсной характеристики или при помощи импульсной характеристики и отдельно измеренного фонового шума (промежуточное измерение). Это значение не всегда совпадает с наблюдаемым отношением сигнал/шум, используемым в вычислениях разборчивости речи. В некоторых сценариях измерения, описанных далее, измеренный уровень сигнала изменяется с учетом упомянутого ранее спектра речи.

Время затухания ранних отражений (EDT)

Так как время затухания ранних отражений EDT (от англ. Early Decay Time) больше относится к начальному участку затухания, характеризующемуся большей энергией, оно больше зависит и от снижения модуляции. Время EDT вычисляется из характеристики затухания и определяется на уровне, когда сигнал падает ниже начального значения от 0 до 10 дБ. На соответствующем склоне значение EDT вычисляется как время достижения сигналом уровня -60 дБ.

Измерение разборчивости речи

Способы измерения

Существуют два фактора, определяющих разборчивость речи. Фоновый шум, вернее, отношение сигнал/шум, и акустика помещения (например, наличие в нем отражений). На стр. 16 влияние отношения сигнал/шум на разборчивость речи рассмотрено более подробно, а также показан важный критерий (отношение сигнал/шум 15 дБ):

фоновым шумом можно пренебречь, если отношение сигнал/шум превышает значение 15 дБ для каждого октавного диапазона, а индекс STI не превышает значения 0,8.

На практике индекс STI редко становится больше 0,8, поэтому в большинстве практических случаев достаточно выполнения одного условия ($SNR > 15$ дБ). Это значит, что измерение фонового шума не требуется, а индекс STI может быть вычислен в любом измерении импульсной характеристики. В большинстве интересующих нас случаев отношение сигнал/шум лежит ниже 15 дБ, поэтому уровень фонового шума важен. В оставшейся части этого документа подразумевается, что критерий 15 дБ не выполняется, если не указано обратное.

В ПО DIRAC акустические параметры помещения вычисляются из измеренной импульсной характеристики. Райф (Rife) [4] определил, что (в определенных условиях) отношение сигнал/шум также может вычисляться по измеренной импульсной характеристике. Это означает, что оба фактора, влияющие на разборчивость речи, могут быть измерены по одному принципу. Условия, требующиеся для того, чтобы определение отношения сигнал/шум по импульсной характеристике было точным: низкий уровень искажений в измерительной цепи, инвариантность акустических свойств ко времени (отсутствует движение воздуха и перемещающиеся люди) и создаваемый системой уровень шума много меньше фонового. Кроме того, нельзя использовать предварительное усреднение, так как оно приводит к искусственно завышенному отношению сигнал/шум.

Следует также учесть, что оба фактора могут быть измерены независимо друг от друга. В самом деле, измерение акустики помещения без фонового шума (при низком его уровне) при помощи импульсной характеристики дает намного лучшие результаты, так как можно использовать предварительное усреднение, последовательности более длинных импульсов и т. п., которые позволяют улучшить качество (отношение амплитуды импульса к шуму (INR)) импульсной характеристики. Также не всегда возможно или не всегда желательно измерять импульсную характеристику при наличии фонового шума.

Измерить фоновый шум проще (и более точно) путем непосредственного измерения, а не косвенным методом при помощи импульсной характеристики.

Источники звука

Источник звука в виде громкоговорителя "искусственный рот"

Направленный источник звука искусственный рот выполнен в виде имитатора рта и имеет диаграмму направленности, аналогичную человеческому рту. Эта диаграмма направленности пригодна для определения разборчивости речи, так как сигнал речи максимален по оси направления источника. Предпочтение следует отдавать искусственному рту с диаграммой направленности, соответствующей ITU-T P.51 [5]. Пригодны также небольшие высококачественные громкоговорители, диаметр которых не превышает 100 мм.

Имитатор рта используется, как правило, в тех случаях, когда следует создать условия, соответствующие речи оратора без усиления, а также тогда, когда система оборудована близко расположенным ко рту микрофоном.

Источник речи с эхом (Echo Speech Source) модели 4720 специально предназначен для измерения раз-

борчивости речи в ПО DIRAC. Он обладает нужной диаграммой направленности, а его сигнал полностью откалиброван как по амплитуде, так и по спектру. Применение этого прибора (Echo) в ПО DIRAC описывается далее в этом документе.

Источник звука в виде всенаправленного громкоговорителя

Всенаправленный источник звука необходим для измерения в соответствии с требованиями стандарта ISO 3382-3. Источник звука должен отвечать требованиям стандарта ISO 3382-1 и должен быть применен на высоте 1,2 м. Всенаправленные источники звука OmniSource модели 4295 и OmniPower модели 4292-L могут использоваться для измерения в офисах с открытой планировкой.

Возбуждение

Изменение частоты или последовательность импульсов максимальной длины

Для измерения импульсной характеристики может использоваться любой широкополосный сигнал, однако, чаще всего используется сигнал с изменением частоты (sweep) и последовательность импульсов максимальной длины (MLS). Экспоненциальное изменение частоты (e-sweep) обладает рядом практических свойств, делающих его очень удобным для измерения импульсной характеристики [9]. Недостатком сигнала с экспоненциальным изменением частоты (или просто с изменением частоты) является то, что такой звуковой сигнал более навязчив по сравнению с последовательностью импульсов максимальной длины. Существуют также условия, в которых сигнал с изменением частоты просто невозможен. Кроме того, изменение частоты делает работу с источником звука более сложной. Последовательность импульсов максимальной длины более удобна на практике и является альтернативой сигналу с изменением частоты.

Прерывающиеся сигналы возбуждения

Чтобы выполнить независимое измерение импульсной характеристики IR (от англ. impulse response) и фонового шума, в ПО DIRAC создается прерывающийся сигнал возбуждения. Такой сигнал состоит из стандартной детерминированной последовательности импульсов максимальной длины или из сигнала с изменением частоты, за которым следует период тишины. Обе части сигнала обрабатываются как отдельное измерение, а полученная в результате импульсная характеристика и уровень фонового шума объединяются в одном wav-файле, в котором импульсная характеристика расположена в одном канале, а фоновый шум – в другом. Прерывающийся сигнал возбуждения также не подвержен описанному Райфом условию, касающегося искусственного увеличения отношения сигнал/шум в случае предварительного усреднения.

Уровень речи и спектр

Измерение разборчивости речи должно выполняться при помощи тестовых сигналов, которые аналогичны человеческой речи по амплитуде и спектру. Это позволяет правильно вычислить отношение сигнал/шум для каждого диапазона частот. Стандарт МЭК 60268-16 определяет спектр мужского и женского голоса для измерения индекса STI, STIPA и STITEL. Параметр STIPA измеряется только для спектра мужского голоса. А-взвешенный уровень сигнала речи должен быть равен 60 дБ на расстоянии 1 м от источника, а октавные диапазоны должны иметь относительные уровни, как показано в табл. 7.

Отдельное измерение импульсной характеристики и фонового шума является предпочтительным методом измерения разборчивости речи в ПО DIRAC. Этот метод может быть совмещен с вычислением разборчивости речи. Кроме того, он делает возможным использовать измеренную отдельно импульсную характеристику для изучения различных сценариев при различных отношениях сигнал/шум.

Таблица 7.
Относительные уровни в октавных диапазонах, дБ

Октавный диапазон, Гц	125	250	500	1000	2000	4000	8000
Мужской	2,9	2,9	-0,8	-6,8	-12,8	-18,8	-24,8
Женский	–	5,3	-1,9	-9,1	-15,8	-16,7	-18,0

Для измерения индекса RASTI, в котором используются только два октавных диапазона, относительные уровни представлены в табл. 8.

Таблица 8.
Относительные уровни в октавных диапазонах для испытания RASTI

Октавный диапазон, Гц	500	2000
Уровень, дБ	-1,0	-10,0

Стандарт ISO 3382-3 использует спектр, вычисленный путем усреднения мужского и женского голоса [10]. Полное А-взвешенное значение амплитуды составляет 57,4 дБ на расстоянии 1 м от (всенаправленного) источника звука, уровень сигнала в каждой октаве которого представлен в табл. 9.

Таблица 9.
Уровень звукового давления речи в соотв. со стандартом ISO 3382-3

Октавный диапазон, Гц	125	250	500	1000	2000	4000	8000
Уровень всенаправленного источника	49,9	54,3	58,0	52,0	44,8	38,8	33,5

Несмотря на то, что ПО DIRAC позволяет создать сигнал с нужным спектром, в большинстве измерительных сценариев на практике применяется произвольный, но обладающий известным спектром источник сигнала, после чего измеренные отношения сигнал/шум вычисляются заново для требуемого спектра речи.

Заметим, что при генерации сигнала с заранее известным спектром должна учитываться частотная характеристика источника. Несмотря на то, что этот принцип может быть осуществлен при помощи эква-

лайзера или фильтра с изменяемой формой, созданного в ПО DIRAC, значительно более просто использовать откалиброванный источник звука.

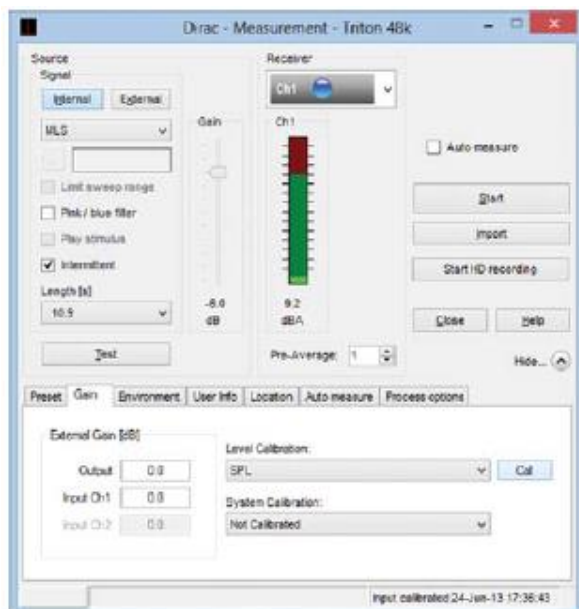
Когда сигнал инжектируется непосредственно в существующую систему оповещения, эквализация не требуется, так как характеристика системы оповещения всегда воздействует на сигнал речи, который является частью системы, подлежащей измерению. Заметим также, что если критерий 15 дБ не выполняется, сигнал возбуждения должен иметь форму речи.

Калибровка уровня

Чтобы измерить абсолютный уровень звука в ПО DIRAC, необходимо выполнить калибровку входного сигнала. Сделать это можно следующим образом:

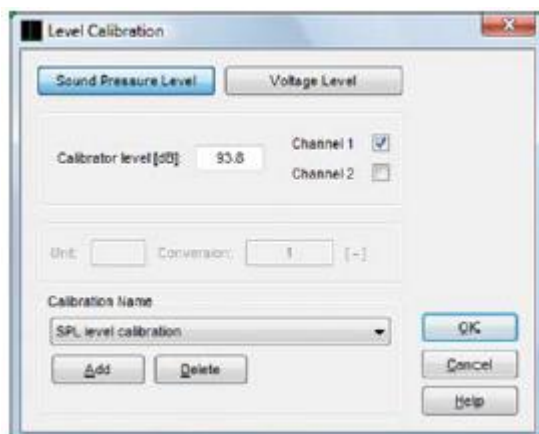
1. нажать  кнопку **Measure** (Измерение) и выбрать вкладку Gain (Усиление), расположенную в нижней части измерительного окна;

Рисунок 2.
Калибровка уровня



2. подключить всенаправленный микрофон к входу звуковой карты, которую предполагается использовать;
3. вставить микрофон в звуковой калибратор одного тока и включить калибратор;
4. нажать кнопку **Cal** (Калибровка) и дождаться окончания измерения. Убедиться, что сигнал калибровки не прекращается до конца измерения;
5. в диалоговом окне Level Calibration (Калибровка уровня) ввести откалиброванный уровень звука, предоставленный производителем, выбрать канал, через который осуществлялась запись калибровочного тона, а затем нажать кнопку **OK**. Теперь уровень входного сигнала для выбранного канала откалиброван;

Рисунок 3.
Диалоговое окно калибровки уровня



6. проверить калибровку, снова включив калибратор, и считать показания (в дБ), расположенные под измерителем уровня звука на экране. Они должны быть равны уровню калибровки.

Примечание:

- Последняя дата калибровки показана в нижнем правом углу окна Measurement (Измерение).
- Пользователь может сбросить результаты калибровки, нажав кнопку **Clear** (Очистить). При этом стираются также записи о калибровке в окне Measurement (Измерение).
- Если с момента сброса последней калибровки канал 1 был откалиброван, а канал 2 нет, канал 2 будет адаптирован к результатам канала 1. Это не даст такие же точные результаты, как если бы канал 2 калибровался отдельно (что рекомендуется), однако при работе с микрофонами одного типа это

все же лучше, чем не калибровать канал 2 совсем. Канал 1 ни при каких обстоятельствах не адаптируется к результатам канала 2.

- Калибровка уровня входного сигнала действительна только для того микрофона и для той звуковой карты, для которых она выполнялась.
- Процедура описана для одного микрофона. Пользователь может также откалибровать два микрофона одновременно, используя калибратор интенсивности звука.
- Калибровка уровня входного сигнала полезна только в том случае, если частотная характеристика звуковой карты имеет плоский участок (± 1 дБ) в диапазоне от 88 Гц до 11,3 кГц.
- Рекомендуется повторять калибровку уровня входного сигнала перед каждой сессией измерения разборчивости речи.

Калибровка системы

Рисунок 4.
Диалоговое окно
калибровки си-
стемы



Источник звука должен создавать сигнал речи с заданным уровнем звука. В большинстве измерений разборчивости речи применяется источник Echo Speech Source модели 4720. ПО DIRAC позволяет распознавать (полностью откалиброванные) сигналы прибора Echo, что делает сочетание Echo-DIRAC самым удобным для точного измерения разборчивости речи.

Измерение на базе стандарта ISO 3382-3 требует применения всенаправленного источника звука, который должен быть откалиброван. ПО DIRAC позволяет сделать это при помощи калибровки системы. Калибровка системы характеризуется полной калибровкой измерительной цепи, в которую входят усилители и микрофоны. Лучше всего выполнять калибровку в помещении с реверберацией, однако, возможна калибровка и в свободном звуковом поле.

Калибровка системы выполняется следующим образом:

1. настроить окно измерения так, как если бы выполнялось измерение разборчивости речи. Убедиться, что выбранный сигнал возбуждения и приемник одного типа, однако, прерывающийся сигнал возбуждения выбирать не следует. Например, если требуется выполнить измерение по стандарту ISO 3382-3, нужно загрузить предустановленные настройки для измерения ISO 3382-3, но выключить в нем функцию Intermittent (Прерывающийся сигнал);
2. запомнить коэффициент усиления, выставленный на усилителе (усилителях). Эта настройка должна оставаться неизменной до конца измерения, в противном случае ее изменение должно быть введено на вкладке Gain (Усиление);
3. выполнить измерение импульсной характеристики для камеры с реверберацией или в свободном звуковом поле (например, на большой сцене). В камере с реверберацией измерение выполняется как минимум в трех точках расположения микрофона. В свободном звуковом поле следует поворачивать источник на 5, 7 или 9 углов от окружности ($360/n$);
4. в меню **Setup** (Настройка) выбрать пункт **Calibrate System...** (Калибровка системы);
5. нажать кнопку **Add** (Добавить) и выбрать калибровочные измерения;
6. выбрать Diffuse field (Диффузное поле) и ввести объем (Volume) камеры с реверберацией или же выбрать Free field (Свободное поле) и ввести соответствующее время окна (Time window), чтобы удалить отражения из сигнала;
7. ввести соответствующее имя калибровки (Calibration name) или описание и нажать кнопку OK;
8. в измерительном окне на вкладке Gain (Усиление) выбрать созданную калибровку системы до того, как запустить измерение разборчивости речи.

Примечание:

- Пользователь может выполнить калибровку системы после измерения и применить калибровку к измерениям при помощи диалогового окна File Properties (Свойства файла). Окно Project (Проект) позволяет применить калибровку системы к нескольким файлам за одно действие.
- Пользователь может также повторить калибровку системы для каждой конфигурации измерения.
- Временное окно для свободного звукового поля должно быть достаточно малым, чтобы удалить из отклика нежелательные отражения, и при этом достаточно большим, чтобы обеспечить точность измерения на низких частотах. Следует проверить визуально импульсную характеристику на предмет определения компромиссного объема данных.

Измерительные сценарии

Измерения при помощи прибора Echo Speech Source (МЭК 60268-16)

Прибор Echo Speech Source поставляется вместе с программным обеспечением. После его установки на ПК, на котором уже установлена копия ПО DIRAC, программное обеспечение (DIRAC) предоставляет возможность выбрать сигнал Echo в качестве возбуждения и предустановленные настройки для прибора Echo. Выполнить измерение с прибором Echo можно следующим образом:

1. выполнить калибровку уровня микрофона;
2. выбрать настройки Echo на вкладке Preset (Предустановленные настройки) измерительного окна;
3. если прибор Echo расположен перед микрофоном, подключенным к системе оповещения, запустить

сигнал речи и задать усиление системы оповещения таким образом, чтобы уровень речи был нормальным;

4. включить прибор Echo Speech и запустить последовательность импульсов максимальной длины (MLS);
5. в измерительном окне ПО DIRAC нажать кнопку **Start** (Запуск).

ПО DIRAC автоматически распознает присутствие непрерывного или прерывистого сигнала, а также нормальный или повышающийся уровень звука.

Если время реверберации относительно небольшое, а влияние фонового шума значительно, следует использовать прерывистый сигнал. За правило можно считать следующее: прерывистый сигнал используется, если $RT \times SNR < 120$. В противном случае следует использовать непрерывный сигнал. Следует заметить, что для увеличения INR (отношение амплитуды импульса к шуму) при работе с прерывистым возбуждением, можно использовать предварительное усреднение.

Сигнал прибора Echo с повышающимся уровнем должен использоваться только в том случае, если уровень шума очень большой, а предварительное усреднение не позволяет увеличить отношение амплитуды импульса к шуму более чем на 20 дБ от нормального уровня сигнала.

Измерения в офисе с открытой планировкой (ISO 3382-3)

ПО DIRAC имеет предустановленные настройки для измерения по стандарту ISO 3382-3. Процедура для измерения офиса с открытой планировкой выглядит следующим образом:

1. выполнить калибровку системы с измерительным оборудованием (требуется выполнить однократно);
2. выполнить калибровку уровня;
3. на вкладке Preset (Предустановленные настройки) измерительного окна выбрать параметры стандарта ISO 3382-3;
4. в них используется возбуждающий сигнал в виде последовательности импульсов максимальной длины, прошедших через розовый фильтр. Если необходимо, можно использовать сигнал с экспоненциальным изменением частоты (e-sweep), однако калибровку систему в этом случае требуется выполнить для этого сигнала возбуждения;
5. при помощи кнопки Test (Тест) определить нужное выходное усиление. Изменить коэффициент усиления усилителя, если это необходимо, и ввести изменение усиления на вкладке Gain (Усиление);
6. в измерительном окне ПО DIRAC нажать кнопку **Start** (Запуск);
7. повторить измерения в 6-10 точках расположения рабочей станции, которые должны находиться на одной линии;
8. выполнить анализ импульсной характеристики.

Предустановленные настройки для измерения по стандарту ISO 3382-3 используют прерывистый сигнал возбуждения. Чтобы увеличить отношение амплитуды импульса к шуму (INR) в измерении, можно задать более высокое количество предварительных усреднений.

Измерение разборчивости речи в других случаях

Если критерий 15 дБ выполняется (см. методы измерения), измерение разборчивости речи может быть выполнено следующим образом:

1. подключить сигнал возбуждения к имитатору рта, расположенному в месте расположения оратора, или к входу звуковой системы;
2. подключить всенаправленный микрофон к входу звуковой карты 1 и установить его в месте расположения слушателя;
3. в ПО DIRAC выбрать сигнал возбуждения (как правило, это экспоненциальное изменение частоты или последовательность импульсов максимальной длины);
4. установить время захвата сигнала более 2 с и оценить время реверберации. Следует заметить, что более длительное время захвата обеспечивается при лучшем показателе INR (отношение амплитуды импульса к шуму);
5. задать количество предварительных усреднений (Pre-average). Более высокое количество предварительных усреднений увеличивает отношение амплитуды импульса к шуму (INR);
6. измерить и сохранить импульсную характеристику;
7. если это возможно, повторить измерение с расположением приемника в другом месте;
8. выполнить анализ импульсной характеристики.

Если критерий 15 дБ не выполняется, измерение импульсной характеристики в условиях наличия фонового шума с целью определения разборчивости речи может быть затруднено, если отсутствует откалиброванный источник сигнала. Чтобы достоверно вычислить отношение сигнал/шум в ПО DIRAC, необходимо, чтобы отношение амплитуды импульса к шуму (INR) было больше 20 дБ. В то же время, низкое отношение сигнал/шум (SNR) в процессе измерения определяет качество импульсной характеристики и, следовательно, отношение амплитуды импульса к шуму (INR). Проблема усложняется тем, что метод вычисления отношения сигнал/шум из импульсной характеристики, разработанный Райфом, не дает достоверных результатов, когда отношение сигнал/шум увеличивается искусственно при помощи предварительного усреднения:

1. откалибровать уровень входного сигнала;
2. при помощи эквалайзера и розового шума сделать характеристику имитатора рта плоской, если это возможно. В ПО DIRAC для этого можно использовать также конструктор формы фильтра (Shaping Filter Designer – SFD), который позволяет изменить форму сигнала возбуждения для определенного источника звука;
3. создать последовательность импульсов максимальной длины (MLS) или сигнал возбуждения с линейным изменением частоты (Lin-Sweep) (меню Edit (Правка), пункт Generate (Генерация)) и отфильтровать его при помощи нужного фильтра (меню Edit (Правка), пункт Filter (Фильтр)). Полученный в результате сигнал следует использовать в качестве заданного пользователем сигнала возбуждения.

- Выбор фильтра зависит от параметров, которые должны быть вычислены (см. табл. 10);
4. задать выходной уровень, равным конкретному значению, или равным 60 дБ (А);

измерить импульсную характеристику (IR) с количеством предварительных усреднений, равным 1, в режиме непрерывного измерения. Увеличить количество усреднений, если это требуется, можно, в режиме прерывистых измерений.

Таблица 10.
Выбор выходного фильтра в зависимости от параметра разборчивости речи

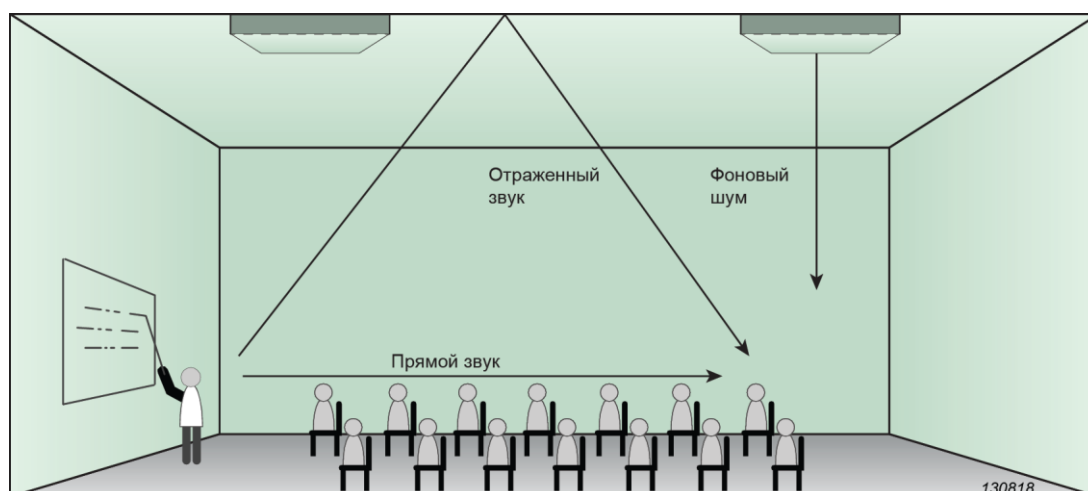
Измеряемый параметр разборчивости речи	Фильтр
STI мужского голоса или соответствующее значение % ALC	Мужской (Male)
STI женского голоса или соответствующее значение % ALC	Женский (Female)
RASTI или соответствующее значение % ALC	RASTI
STIPA или соответствующее значение % ALC	Мужской (Male)

Заметим, что можно использовать как созданный в программе (внутренний), так и при помощи оборудования (внешний) сигнал возбуждения. При работе с внешним возбуждением, тип возбуждения, фильтр источника и длительность последовательности этого сигнала должны соответствовать настройкам измерительного диалогового окна. Ряд сигналов в виде отфильтрованной последовательности максимальной длины (MLS) и сигналов с линейным изменением частоты (Lin-Sweep), предназначенных для измерения разборчивости речи, находятся на компакт-диске с ПО DIRAC.

На этом этапе важно убедиться, что отношение амплитуды импульса к шуму (INR) для всех значимых октавных диапазонов измеренной импульсной характеристики имеет значение более 20 дБ. Это необходимо в ПО DIRAC для вычисления отношения сигнал/шум и многих других параметров. В большинстве случаев это требование можно легко выполнить, комбинируя различные методы, например, увеличивая длительность захвата сигнала, количество предварительных усреднений и уровень выходного сигнала возбуждения.

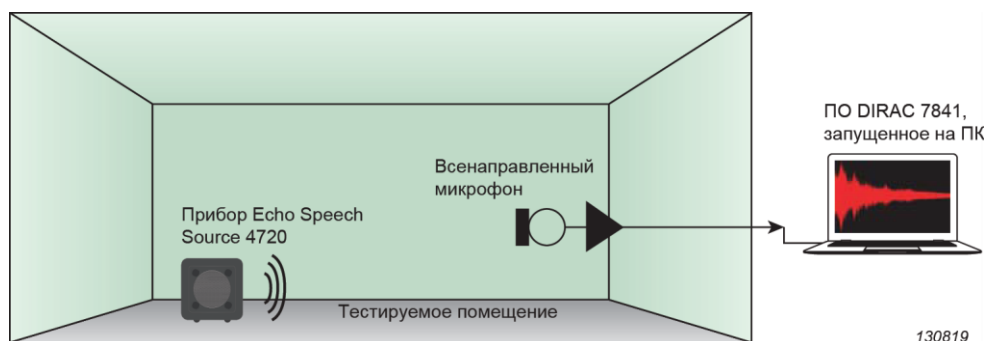
Ситуация 1: оратор и слушатели находятся в одном помещении без звуковой системы

Рисунок 5.
Помещение без звуковой системы



Слушатели принимают звук непосредственно от оратора, отраженный звук, такой как реверберация и эхо, и фоновый шум, например, от системы вентиляции и охлаждения (HVAC). Вклад прямого звука позитивно влияет на разборчивость речи. Отраженный звук может влиять на разборчивость речи как позитивно, например, если он отражен от передней стенки, так и негативно. Фоновый шум негативно сказывается на разборчивости речи.

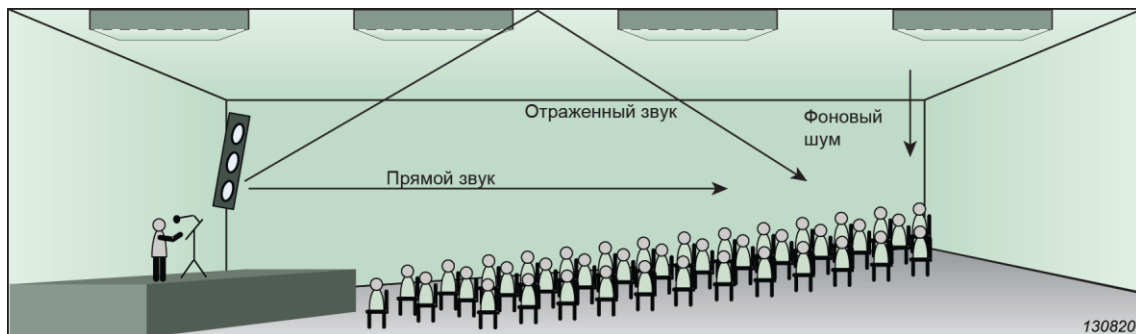
Рисунок 6.
Измерительная установка в помещении без звуковой системы



Ситуация 2: оратор и слушатели находятся в одном помещении со звуковой системой

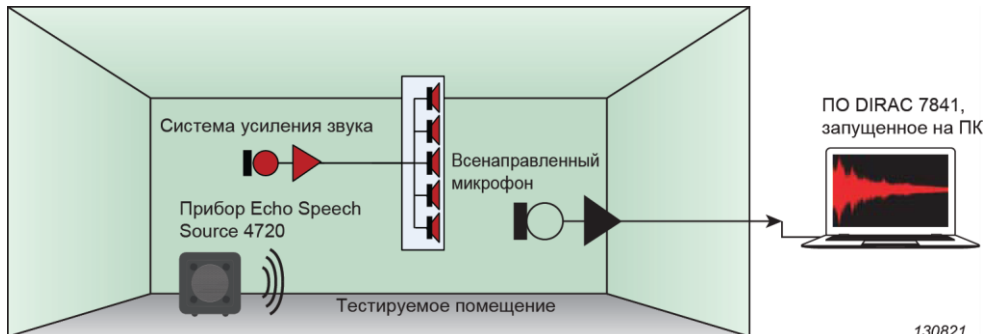
На рис. 7 показано помещение с системой усиления звука, в котором находятся оратор и слушатели. Слушатели принимают звук, главным образом, из системы усиления звука, а также отраженный звук, такой как реверберация и эхо, и фоновый шум, например, от системы вентиляции и охлаждения (HVAC). На это накладывается ряд некоторых побочных эффектов. Прямой звук, идущий от оратора, обладает низкой амплитудой, является задержанной версией основного прямого звука и тоже доходит до слушателей. Следовательно, микрофон звуковой системы улавливает не только прямой звук, идущий от оратора, но и отражения от кафедры, аналогичный обратной связи звук и фоновый шум.

Рисунок 7.
Помещение с системой усиления звука



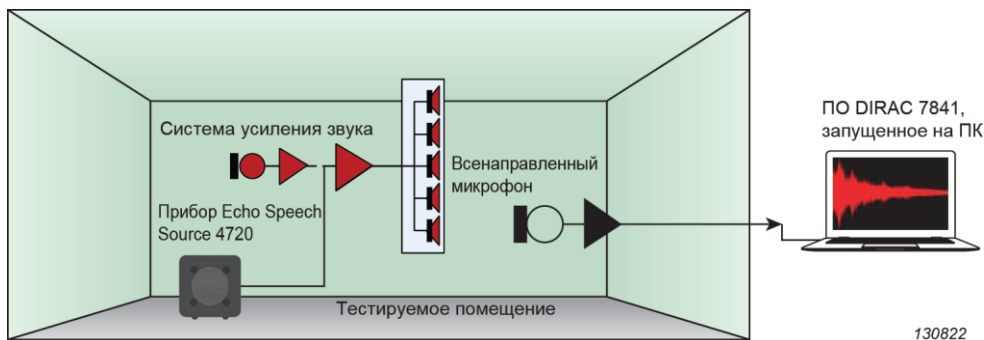
В этой ситуации существует два способа измерения разборчивости речи. Чтобы учесть характеристики микрофона звуковой системы и все побочные эффекты, необходимо работать с имитатором рта.

Рисунок 8.
Измерения в помещении со звуковой системой



Также сигнал возбуждения может быть инжектирован непосредственно в звуковую систему.

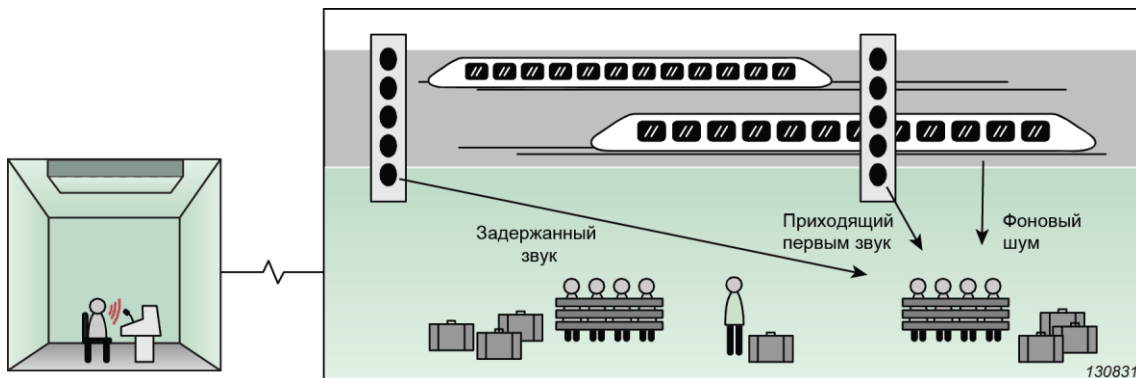
Рисунок 9.
Измерительная установка с прямой инъекцией для помещения со звуковой системой



Ситуация 3: оратор и слушатели расположены в разных областях

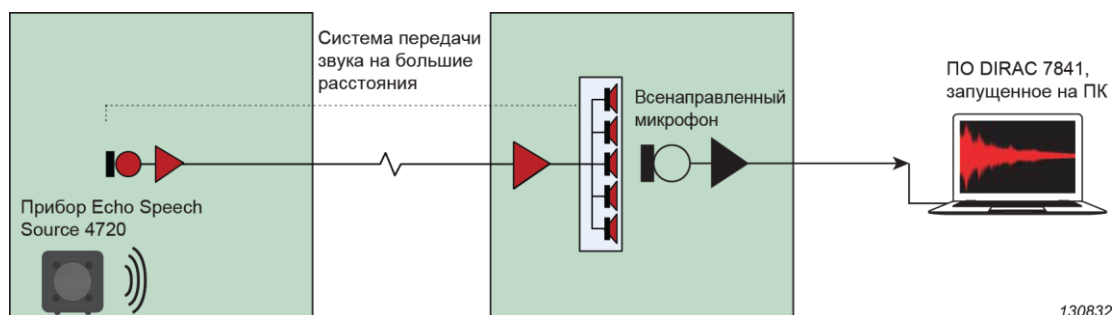
На рис. 10 показан случай, когда оратор и слушатели расположены в отдельных областях, связанных между собой протяженной звуковой системой. Большей частью слушатели слышат прямой звук, исходящий от расположенного рядом массива громкоговорителей, задержанный прямой звук, исходящий из массива громкоговорителей, расположенных на удалении, а также реверберацию или эхо и фоновый шум, например, окружающего пространства. На это накладывается ряд некоторых побочных эффектов. Микрофон звуковой системы улавливает не только прямой звук, идущий от оратора, но и отражения, и фоновый шум, присутствующие в комнате с оратором.

Рисунок 10.
Оратор и слушатели в разных помещениях, объединенных протяженной звуковой системой



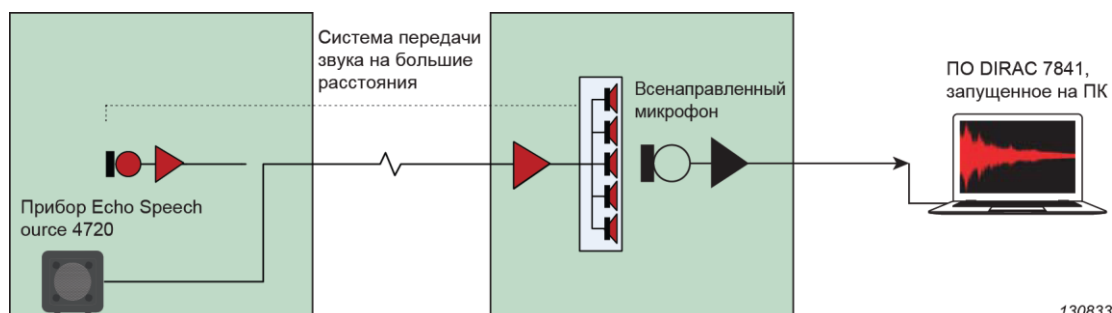
В этой ситуации также существует два способа измерения разборчивости речи. Чтобы учесть характеристики микрофона звуковой системы и все побочные эффекты, необходимо работать с имитатором рта.

Рисунок 11.
Измерительная установка с разомкнутой петлей обратной связи и прибором Echo Speech



130832

Рисунок 12.
Измерительная установка с разомкнутой петлей обратной связи, использующая прямую инъекцию сигнала



130833

Анализ измерения

Качество импульсной характеристики

Проверять качество измеренной импульсной характеристики необходимо после каждого измерения, так как характеристика низкого качества может стать причиной ошибочных результатов. Существует три свойства импульсной характеристики, которые могут быть легко проверены и которые дают хорошее представление о качестве и пригодности импульсной характеристики.

Первым из них является отношение амплитуды импульса к шуму (INR – от англ. Impulse to Noise Ratio). Величина INR может быть интерпретирована как диапазон спада импульсной характеристики. В общем случае, чем INR больше, тем лучше. Минимальной приемлемой величиной является значение 20 дБ. Величина более 40 дБ является предпочтительной. Необходимо заметить, что значение INR может быть увеличено за счет сигнала большей амплитуды, за счет увеличения времени захвата, а также использованием предварительного усреднения.

Другим свойством является точка корректного обнаружения фронта импульсной характеристики в ПО DIRAC. Начало импульсной характеристики в окне просмотра отмечается вертикальным маркером красного цвета.

Рисунок 13.
Слева: правильное обнаружение фронта импульсной характеристики. Справа: ложный пик, вызывавший ошибочное обнаружение.



Последним свойством импульсной характеристики, которое следует проверить, является следующее: достаточно ли длительность захвата сигнала для данной импульсной характеристики? За правило можно взять следующее: длительность захвата должна равняться двум временам реверберации. Это свойство можно проверить другим способом: точка, в которой импульсная характеристика начинает тонуть в фоновом шуме, должна находиться в пределах измерительного окна. Окно измерения должно быть таким, чтобы плоский хвост импульсной характеристики, переходящий в фоновый шум, был представлен на экране.

Таблица параметров и графики

Таблица Speech (Речь), которая может быть открыта при помощи меню Parameters (Параметры) окна Impulse Response (Импульсная характеристика), содержит в себе список параметров разборчивости речи для данной характеристики.

Рисунок 14.
Таблица Speech
(Речь) с вычисленными параметрами речи

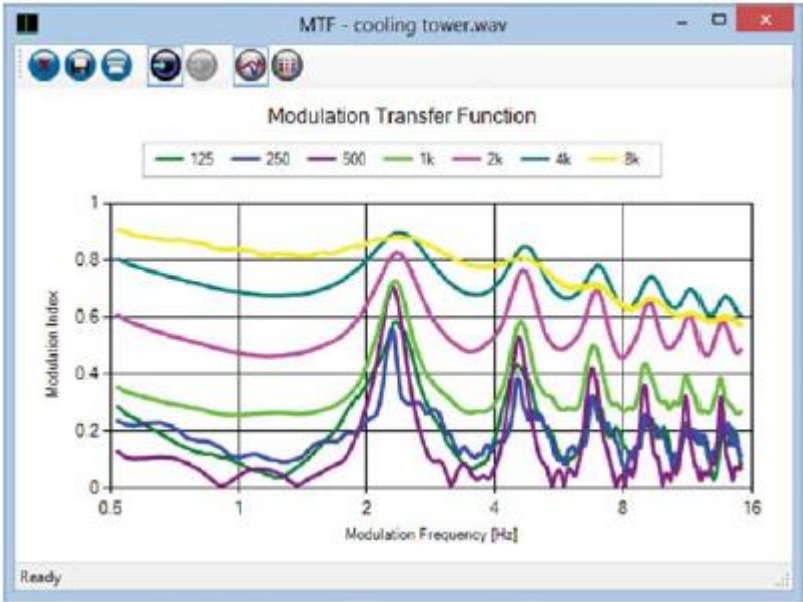


Необходимо заметить, что отношение сигнал/шум (SNR) в этой таблице показывает величину, измеренную в точке расположения микрофона. Для вычисления разборчивости речи отношение сигнал/шум не обязательно, так как ПО DIRAC может изменить уровень сигнала в соответствии с определенным спектром речи таким образом, чтобы нормализовать источник звука.

Таблица с параметрами речи позволяет быстро получить доступ к свойствам файла, если нажать кнопку , кнопка Properties (Свойства), где можно изменить уровень сигнала и уровень шума (см. стр. 16).

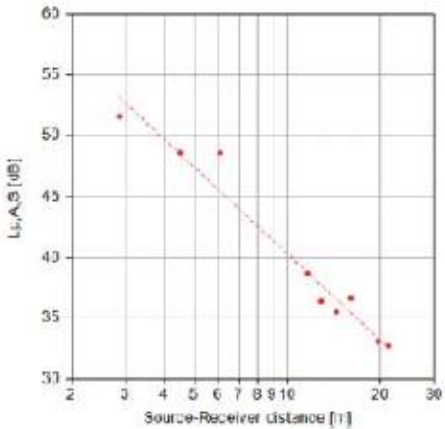
График функции MTF выводится на экран нажатием кнопки MTF, , которая позволяет изучить причину снижения модуляции (см. стр. 3).

Рисунок 15.
График MTF для охлаждающей камеры



Необходимо заметить, что график MTF, описанный выше, может быть вычислен только в измерении импульсной характеристики, так как в измерении модулированным шумом содержится только ограниченный набор дискретных модулированных частот.

Рисунок 16.
Линия регрессии, проходящая через несколько измерений



Отдельные параметры и статистические значения по нескольким файлам могут быть выведены на экран в окне Parameter Graph (График параметров). На нем показано среднее, минимальное и максимальное значение, а также их среднеквадратическое отклонение. По этим графикам можно также построить зависимость параметров от расстояния источник-приемник, а также вычислить параметры, базирующиеся на линии регрессии, такие как $L_p, A, S, 4$, $D2, S$, rD и rP , определенные в стандарте ISO 3382-3.

Изменение уровня сигнала и уровня шума

Часто требуется знать, какая разборчивость речи будет, если уровень сигнала или шума будет отличаться от тех, которые были в момент измерения импульсной характеристики. ПО DIRAC предоставляет несколько способов исследования результата измерения при изменении уровня речи или уровня шума.

Когда импульсная характеристика измеряется в присутствии небольшого фоновых шума, существует возможность добавить шум в измерение. Это может быть синтетический шум (пункт меню: Process (Обработка) – Mix (Смешивание) – Noise (Шум)), или запись реального шума (пункт меню : Process (Обработка) – Mix (Смешивание) – File (Файл)).

В окне File Properties (Свойства файла) можно задать уровень шума и сигнала речи при помощи числовых значений. Чтобы исследовать эффект увеличения громкости речи или уровня шума, в качестве отправной точки можно использовать значение параметра Speech (Речь) и Noise (Шум).

Рисунок 17.
Изменение уровня шума



Заметим, что можно ввести как уровень шума, так и уровень речи, а также оба этих значения одновременно. ПО DIRAC будет использовать то значение, которое было введено в диалоговом окне File Properties (Свойства файла), а в тех случаях, когда данных введено не было, будет использовать реальные значения, полученные путем измерения.

Окно графика параметров позволяет редактировать свойства сразу для нескольких файлов. Меню Edit (Правка) позволяет изменить свойства всех файлов, которые были отмечены.

Влияние отношения сигнал/шум на разборчивость речи

Если уровень фоновых шума пренебрежимо мал, измерять отношение сигнал/шум и учитывать его при вычислении разборчивости речи не требуется. Поэтому может быть полезно более детально разобрать влияние отношения сигнал/шум на разборчивость речи в практическом случае, а также сформулировать точные условия, в которых его следует учитывать.

Фоновым шумом можно пренебречь, если в его присутствии индекс STI снижается не более чем на 5 % по сравнению с индексом STI, измеренным без фоновых шума. Чтобы выразить влияние отношения сигнал/шум, запишем функцию MTF как произведение двух коэффициентов снижения модуляции, $m_0(k, F)$, образованный свойствами системы (реверберация, эхо) и $m_{SNR}(k, F)$, образованный фоновым шумом:

$$MTF(k, F) = m_0(k, F) \cdot m_{SNR}(k) = \frac{m_0(k, F)}{1 + 10^{-\frac{SNR(k)}{10}}}$$

где k – номер октавы. По определению, эффективное отношение сигнал/шум, $SNR_{eff}(k, F)$ относится к MTF(k, F) как SNR к m_{SNR} :

$$SNR_{eff}(k, F) = 10 \log \left(\frac{MTF(k, F)}{1 - MTF(k, F)} \right)$$

Это значит, что SNR_{eff} равняется SNR, если модуляция уменьшена только за счет фоновых шума. Для каждого k из набора октавных диапазонов и для каждого F из набора частот модуляции вычисляется величина $SNR_{eff}(k, F)$, которая ограничивается на уровне ± 15 дБ до того, как приступить к дальнейшему вы-

числению индекса STI. Операция ограничения отражает тот факт, что величина SNR_{eff} , превышающая 15 дБ, не оказывает негативного влияния на разборчивость речи, в то время как величина SNR_{eff} менее минус 15 дБ не способна оказать положительного воздействия на нее. Ограниченные значения SNR_{eff} преобразуются в индексы передачи $TI(k, F)$, которые, принимая значение от 0 до 1, вносят свой вклад в индекс STI.

$$TI(k, F) = \frac{SNR(k, F) + 15}{30}$$

Конечное значение STI получается путем усреднения особым образом индексов передачи TI по всем частотам модуляции и октавным диапазонам, что позволяет учесть маскирование аудитории и абсолютный порог слышимости. Тем не менее, чтобы понять смысл влияния отношения сигнал шум (SNR) на индекс передачи речи (STI), достаточно оценить индекс передачи (TI) для нескольких значений m_0 и SNR. На рис. 19 показано относительное изменение индекса TI в зависимости от ситуации, когда $SNR = \infty$ и когда SNR имеет конечное значение:

$$\Delta TI = \frac{TI_{SNR} - TI_{\infty}}{TI_{\infty}} \cdot 100\%$$

Относительные изменения индексов TI сохраняются в индексе STI при условии равных весовых коэффициентов для каждого октавного диапазона частот. На самом деле, наиболее важными являются значения SNR для октавных диапазонов с 500 Гц по 4 кГц.

На рис. 18 и 19 видно, что при увеличении m_0 более 0,5 изменение индекса TI происходит быстрее (а значит, и отношения сигнал/шум (SNR)) до тех пор, пока значение SNR_{eff} не будет ограничено на уровне 15 дБ при $TI = 1$, в результате этого образуется провал при $m_0 = 0,97$. Еще больший провал наблюдается при $m_0 = 0,03$, но так как он не встречается на практике, он не показан.

Используя рис. 19 и табл. 10, в которых дается определение относительным октавным диапазонам частот, мы можем сформулировать критерий для **отношения сигнал/шум, равного 15 дБ**:

фоновым шумом можно пренебречь, если отношение сигнал/шум превышает значение 15 дБ для каждого октавного диапазона, а индекс STI не превышает значения 0,8.

Если выполняется только одно условие: отношение сигнал шум более 15 дБ, это, теоретически, может привести к недооценке индекса STI на 10%. Другими словами, если SNR лежит в районе 15 дБ для всех октавных диапазонов, и индекс STI без фонового шума равен 0,97, результат измерения индекса STI даст значение 0,9. Однако на практике индекс STI редко превышает значения 0,8. Гораздо более вероятно, что только несколько значений SNR будут близки к порогу 15 дБ. Поэтому с практической точки зрения достаточно выполнения только одного условия: отношение сигнал шум (SNR) должно быть больше 15 дБ.

Рисунок 18.
Индекс передачи как функция от индекса модуляции при отсутствии фонового шума

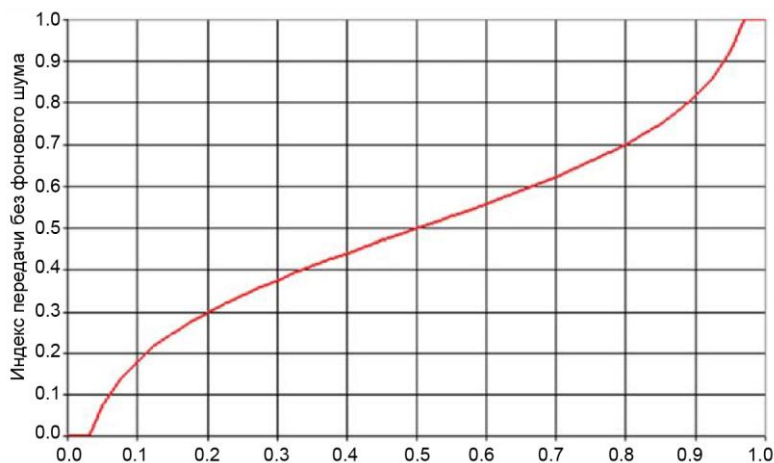


Рисунок 19.
Относительное изменение индекса передачи как функция от индекса модуляции для различных уровней фонового шума

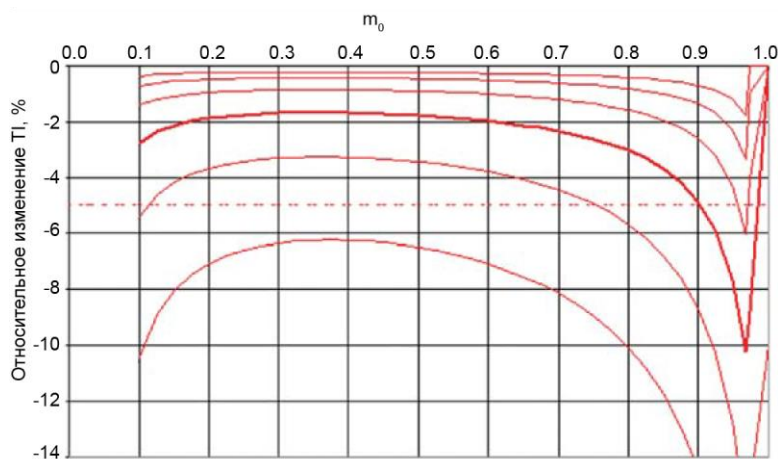


Таблица 11.
Относительные октавные диапазоны

Октавный диапазон, Гц	Источник сигнала возбуждения		
	Мужской (Male)	Женский (Female)	RASTI
125	●		
250	●	●	
500	●	●	●
1 к	●	●	
2 к	●	●	●
4 к	●	●	
8 к	●	●	

Список литературы

- [1] T. Houtgast, H.J.M. Steeneken and R. Plomp, "Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function. I. General Room Acoustics," *Acustica* 46, 60 – 72 (1980)
- [2] IEC 60268-16 Edition 4.0 (2011) Sound system equipment. Part 16: Objective rating of speech intelligibility by speech transmission index
- [3] M.R. Schroeder, "Modulation Transfer Functions: Definition and Measurement," *Acustica* 49, 179 – 182 (1981)
- [4] D.D. Rife, "Modulation Transfer Function Measurement with Maximum-Length Sequences", *J. Audio Eng. Soc.* 40, 779 – 790 (1992)
- [5] ITU-T Recommendation P.51 (03/93), "Telephone Transmission Quality Objective Measuring Apparatus – Artificial Mouth"
- [6] T. Houtgast, H.J.M. Steeneken et al., "Past, present and future of the Speech Transmission Index", TNO Human Factors, Soesterberg, The Netherlands, © 2002, ISBN 90-76702-02-0
- [7] ISO 3382 (2012), "Acoustics – Measurement of room acoustics parameters – Part 3: Open plan offices"
- [8] "A Do-it-Yourselfers Guide to Computing the Speech Transmission Index", <http://www.synaudcon.com/site/articles/a-do-it-yourselfers-guide-to-computing-the-speech-transmission-index/>
- [9] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique", 110th AES Convention, February 2000
- [10] "Guidance on computer prediction models to calculate the speech transmission index for BB93." Version 1.0, Department for Education and Skills, Schools Capital and Building Division, 2004

