

## «Говорящие» с абонентом Ольга Кузнецова

### Особенности применения IVR и систем распознавания речи



**Услуги систем интерактивных голосовых меню (IVR – Interactive Voice Response), используемые, в том числе в сетях IP-телефонии, прочно вошли в повседневную жизнь миллионов людей. Системы интерактивного голосового взаимодействия нашли свое применение в телекоммуникационной отрасли, офисах крупных корпораций, торговых и сервисных сетях. IVR может произвести обработку телефонного вызова на основе информации, поступающей от пользователя, в том числе осуществить маршрутизацию телефонного соединения в соответствии с пожеланиями пользователя или предоставить доступ к определенной информации.**

До недавнего времени предлагаемые на отечественном рынке системы автоматического информирования обладали одним очень существенным недостатком – они работали только с вводом данных в тональном наборе. Как известно, в нашей стране большая часть телефонных аппаратов – дисковые, они вообще не приспособлены к такому режиму работы. Обладатели кнопочных аппаратов зачастую не пользуются тональным набором из-за того, что он кажется им слишком сложным. К тому же, выслушивая весь перечень предлагаемых вариантов, иногда можно просто «заблудиться» и забыть, какую цифру надо набирать. А если далее следует еще «вложенное» меню, то добраться до цели совсем непросто. «Работать» с клавишами телефона достаточно сложно, особенно в радиотелефонах, когда человек во время разговора вынужден еще и нажимать кнопки на аппарате. Таким образом, есть смысл поговорить о современной технологии обработки речи, которая существенно расширяет возможности пользователей, обращающихся за услугами автоматизированных служб по телефону. Технология распознавания речи позволяет строить приложения телефонных услуг, в которых абонент осуществляет выбор в меню не тональным набором, а произношением отдельных слов или фраз. Данная технология может с успехом использоваться как в системах интерактивного речевого взаимодействия, так и для построения речевых порталов. С помощью этих систем пользователи избавляются от необходимости подолгу оставаться на линии, ожидая ответа оператора. Система распознавания речи позволяет создавать такие решения, которые практически невозможно реализовать, имея в качестве интерфейса только тональный набор. Например, ввод запроса по названию города. А что уж говорить об использовании биометрических технологий, позволяющих установить личность человека по характеристикам его голоса. Таким образом, системы распознавания речи в телефонных сетях открывают принципиально новые возможности для бизнес-приложений на основе IVR-решений. Естественно, что приложения, распознающие живую речь, по сравнению с приложениями, использующими тональный набор, намного сложнее в разработке и внедрении. Именно поэтому разработчик IVR, распознающих живую речь, должен иметь междисциплинарные знания, быть сведущим и в компьютерных технологиях, и в лингвистике, и в психологии. Кроме того, чрезвычайно важны сбор и а вариантах взаимодействия с пользователями. Еще один существенный момент – точность распознавания речи гораздо ниже (по сравнению с распознаванием тонального набора). Здесь приходится учитывать и дополнительные фоновые шумы, и тот факт, что пользователь практически ничем не ограничен при выборе речевых форм. В свою очередь, это повышает степень сложности и увеличение объема программирования, обеспечивающего компенсацию ошибок (низкую точность распознавания, двусмысленность выражений, словарь без грамматики и т. п.). Для реализации телефонного приложения с функциями распознавания речи необходимы: платы компьютерной телефонии со специализированными CSP-ресурсами (Continue Speech Processing), способными выполнять функции непрерывной обработки речи, и программное обеспечение для распознавания произносимых абонентом фраз. Программное обеспечение распознавания речи – это инструментальный для реализации функций распознавания речи в телефонных приложениях с использованием оборудования компьютерной телефонии. Обычно эта платформа основана на таких технологиях, как распознавание речи с большим объемом словарных единиц (Large Vocabulary Recognition), понимание естественного языка (Natural Language Understanding), преобразование текста в речь (Next-to-Speech), а также идентификация говорящего по голосу (Speaker Verification). Основными характеристиками данного ПО являются:

- возможность дикторонезависимого распознавания;

- отсутствие необходимости настройки на голос (однако режим работы с настройкой на голос не исключается);
- распознавание непрерывной речи;
- распознавание сложных предложений, чисел, отдельных фраз и команд;
- поддержка нескольких языков;
- произвольный словарь распознаваемых слов, составляемый разработчиком приложения;
- возможность «перебивать» систему, произнося команды, не дожидаясь окончания проигрывания приветствий и/или меню.

К основным компонентам систем распознавания речи относятся:

- графическая среда для разработки, компиляции и оптимизации грамматических и лексических блоков распознавания, проверки и редактирования лексиконов;
  - система для протоколирования диалогов из работающего приложения в удобном для оценки качества распознавания и подстройки системы виде;
  - инструмент оценки качества работы системы (проверка соответствия слова, сказанного абонентом, используемой грамматике);
  - система, позволяющая создавать «тренируемые» языковые модели, что повышает производительность и ускоряет сам процесс распознавания;
  - система, предназначенная для распределения множества параллельных запросов различных типов для прозрачной интеграции различных речевых модулей в сети.
- Распознавание речи значительно упрощает логику общения клиента с системой IVR. Однако перед внедрением системы с распознаванием речи необходимо определить, будет ли приложение: использовать речь в полном объеме или частично; аварийно переходить на тональный набор. Кроме того, до того, как она закончила, например, подтверждение того, что она поняла).

Хорошая практика – обеспечение аварийного тонового набора как меры предосторожности, например, по соображениям безопасности (при вводе номера кредитной карты, пароля и т. п.), либо если клиент затрудняется в формулировании своих мыслей в виде речи или просто консервативен и боится отказаться от привычного тонального набора.

Одно из важных преимуществ речевых приложений – уменьшение глубины голосового меню, т. е. сокращение количества шагов, которые должен сделать абонент для достижения своей цели. Например, не совершая многоходовых комбинаций, можно проникнуть на нижний уровень меню при помощи одной ключевой фразы. И даже если абонент имеет опыт общения с кнопочным меню, заранее зная, какие кнопки нажимать, все равно доступ к нужной позиции при помощи ключевой фразы значительно экономит время абонента.

Существуют несколько рекомендаций, позволяющих повысить эффективность технологии распознавания речи. Прежде всего, направляющий диалог должен включать вопросы, на которые можно дать однозначный ответ: «да» или «нет». Приветствие необходимо дополнять подсказкой о получении помощи. Абонент должен иметь достаточно времени для передачи запрошенной информации. Следует предусмотреть аварийные ситуации. Если фоновый шум или какие-либо помехи не позволяют программе распознавания речи понять абонента, применимо, к примеру, сообщение: «Извините, я вас совсем не понимаю» или «Мне жаль, но я не расслышала, пожалуйста, говорите громче». В этом случае пользователь получает больше информации от системы, у него не возникает сомнения, что система работает, и он не бросает трубку.

Таким образом, грамотное использование технологии распознавания речи позволяет создать IVR-систему, имитирующую общение с живым человеком.

В настоящее время существует довольно широкий выбор программ преобразования текста в речь. Эти программы преобразуют написанный текст в натурально звучащую речь и