Nuance Communications, Inc.
Who we are
Reinventing the relationship between people and technology

- Defining the next generation of human-computer interaction: Intelligent Systems
- Deeply invested in creating effortless and natural user experiences
- Best known for rapidly advancing voice-recognition technology
We humanize technology

- User centered design
- Game-changing innovations
- Bold strategic moves
  - Pioneering Natural Language Understanding (NLU)
  - Nina
  - Driving the industry from speech to understanding to Intelligent Systems
It’s rare to go a day without Nuance
We’re a company most people don’t realize they already have a relationship with
<table>
<thead>
<tr>
<th>statistic</th>
<th>number</th>
<th>notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 billion mobile cloud</td>
<td>5 billion</td>
<td>transactions annually</td>
</tr>
<tr>
<td>3,900 patents &amp; applications</td>
<td>3,900</td>
<td></td>
</tr>
<tr>
<td>65+ countries</td>
<td>65+</td>
<td></td>
</tr>
<tr>
<td>12 billion customer calls</td>
<td>12 billion</td>
<td>served annually</td>
</tr>
<tr>
<td>12,000 employees</td>
<td>12,000</td>
<td></td>
</tr>
<tr>
<td>10 billion customer</td>
<td>10 billion</td>
<td>service transactions</td>
</tr>
<tr>
<td>70+ languages</td>
<td>70+</td>
<td></td>
</tr>
<tr>
<td>800 million mobile</td>
<td>800 million</td>
<td>keyboards</td>
</tr>
<tr>
<td>10,000 mobile app</td>
<td>10,000</td>
<td>developers</td>
</tr>
<tr>
<td>1,200 voice &amp; language</td>
<td>1,200</td>
<td>scientists &amp; engineers</td>
</tr>
<tr>
<td>5 billion lines of medical transcription</td>
<td>5 billion</td>
<td></td>
</tr>
<tr>
<td>20 million voice-enabled</td>
<td>20 million</td>
<td>cars</td>
</tr>
</tbody>
</table>
What we mean by ‘intelligent systems’

Natural User Interface & Ambient Intelligence
A new world of interaction is opening up

- People expect technology that understands natural input
- As capabilities improve, demand for intelligent automated services continues to rise
- This new realm includes the analytics of human behavior
The ecosystem is quickly perfecting

NLU Innovations

Mixed Modality Designs

Algorithmic Advances

Mobile Device Proliferation

Developer Community

Data Volumes

Computational Resources
Human-centered design is our ‘true north’

- Ease-of-use is the guiding principle
- Everything we do aims for acceptance by customers and cultures
- We solve for every aspect of user experience, using innovation to drive demand for new wants
The Nuance Natural Language Framework
Speech Technology

A little history…
A little history of Speech Technology

- 1938: Bell Labs demonstrates “the Voder”
A little history

- **1952** Bell Labs demonstrates a digit recognizer on an analog computer.

- **1968** The HAL-9000 computer in *2001: A Space Odyssey* introduces the world to speech recognition.

- **1972** DARPA Launches Speech Understanding Research program.

- **1976** IBM, Jim Baker designs a 1,000-word HMM based continuous speech recognizer from a vocabulary.

- **1982** Jim Baker founds Dragon Systems and Ray Kurzweil founds Kurzweil AI.

- **1982** DARPA funds MIT, CMU, SRI, BBN, others.
A little history

- **1990** Dragon Systems introduced DragonDictate 30K
- **1992** Phillips introduces SpeechMagic, the first continuous recognition system for medical transcription
- **1993** IBM launches the IBM Personal Dictation System for OS/2.
- **1993** Apple ships PlainTalk, a series of speech recognition and speech synthesis extensions for the Macintosh.
- **1994** Dragon Systems' DragonDictate for Windows is the first software-only PC-based dictation product.
- **1996** IBM introduces MedSpeak/Radiology, the first real-time continuous-speech recognition product.
- **June 1997** Dragon ships NaturallySpeaking, the first general-purpose continuous-speech recognition product.
- **Fall 1997** Microsoft CEO Bill Gates identifies speech recognition as a key technological advance.
- **2000**: L&H acquires Dragon Systems and Dictaphone
- **October 2000**: L&H in Chapter 11
Launched in April 1997 – first continuous speech recognition, used primarily as assistive technology

Nuance acquired Dragon assets, invested heavily in research and development, boosted Dragon performance on desktop

Dragon speech technology now spans Desktop, Mobile, Automotive, Entertainment. More data = better core accuracy

Today, Dragon enables millions of people to communicate more naturally, and to get more done – by voice
65 Years of R&D – and its not perfect yet?

- Perfect equals on a par with Human understanding
- As viable speech applications have emerged, they have grown in complexity and variability
- Very hard computational problem
- Input Data Volume
- Astronomical Search Space
- Ambiguity
- Variability

- Do what I want, not necessarily what I say!

The Nuance Natural Language Framework gives us the capability to create the next generation of Human-Machine interaction
What does the future hold?
What is the Dragon SDK?

Dragon Developer Edition

Set of COM Interfaces and ActiveX Controls (Methods, Properties and Events)

Speech Engine Runtime (alternatively Dragon Pro can be used)

Tools to Facilitate the Development of Applications (Efenroll, VocTool, ATFA, etc.)

Sample Applications (VB, C++, C#)

Documentation

• API Reference

• Developer’s Guide
Dragon SDK: Use Cases

• Client SDK (DSC)
  • Close Captioning
  • Audio Transcription as part of an application’s feature
  • Creating full dictation support for custom edit control
  • Embedding speech recognition & Command & Control in 3rd party apps enabling specific grammars or vocabularies

• Server SDK (DSS)
  • Speaker Independent recognition – US English models only
  • High quality audio transcription part of a workflow using speaker dependent models (handhelds, smart phones, etc)
Dragon Client SDK
Product Overview

Dragon NaturallySpeaking Engine
Focus on Interactive Dictation Integration into Windows Applications
Control over dictation via compatibility modules
DictEdit and DictCustom controls
Custom command creation
Importing\Exporting commands
Tools for Speaker and Vocabulary Management:
VocTool

Documentation:
API Reference and Developer’s Guide
Set of sample programs

Languages included:
US/UK/AUS/IND/SEA – English, French, Italian, German, Spanish, Dutch
General vocabularies

Acoustic Optimizer (ACO) and Language Model Optimizer (LMO) scheduler
Phrase Hypothesis

Definition

- The Dragon SDK provides an interface called Phrase Hypothesis, which allows applications access the text results as soon as recognition starts.

- SDK interface allows access to recognition results as observed in the yellow results box.

- The engine provides access to unformatted recognition words.

- Needs the use of transformation library (ITN) for numbers.*
Phrase Hypothesis
Interface & Recommendations

- PhraseHypothesis returns the intermediate results of a recognition before the speech engine finishes processing the utterance.

- PhraseHypothesis event is available in DgnDictEdit/DgnDictCustom

- Avoid displaying the microphone VU meter (the status bar indicating the volume of the mic) available in the DgnMicBtn ActiveX control, as that requires additional notifications that can conflict with PhraseHypothesis

- Don't call any Dragon methods or properties in the PhraseHypothesis event, other than the ones required to get the recognized words

- Display or consider final results n-3 words, where 'n' is the number of words received in the event, this because the last 3 words are likely to be changed during the recognition as the engine process more context

- Raw punctuation tokens needs to be processed at application level, by extracting the written form from the token. Ex. “.\period” or “,\comma”
General Performance Recommendations

- Reduce NBest Count (Dropped in V12)*
- Disable confidence scores
- If possible, switch to Dictation-Only mode
- Use noise cancelling Mic
- Disable audio for playback
- Perform Mic calibration (ASW) on every new session
- Intel i7 core CPU gives SR a performance boost
Dragon Client SDK – V12
Available April 2013

- Brings V12 engine accuracy improvements to the SDK

- 2 new improvements for Subtitling applications:

**DgnDictationOptionConstants.dgndictoptionFormatPhraseHypothesis** – This option allows users to enable ITN formatting during a phrase hypothesis callback, this now works for all languages. There’s no longer need to integrate the standalone ITN library that was only supported in English.

**DgnEngineOptionConstants.dgnengoptionNBestParamTopChoiceOnly** – This option if enabled configures the recognizer to output only a single alternative (top choice), this should be used in closed captioning application to speed up recognition performance.
Thank you.