Speech Analytics in Practice
Towards cross-channel real-time awareness

Marsal Gavalda, PhD
Vice President of Incubation and Principal Language Scientist
Outline

Speech analytics in practice
• The science of speech analytics
• Applications of speech analytics

Towards cross-channel real-time awareness in the call center
• Measuring agent performance
• Live agent assist
• Topic discovery and trending
The Science of Speech Analytics
Spoken Content Analysis

Speech to Text

• Audio → transcript
• Search term + transcript → hits
• Need to maintain lexicon and train statistical language model
• High word error rates (~40%) unless enrolled speaker or very constrained task
• Slow (<1 to 4 xRT, with severe speed vs. accuracy tradeoff)

Phonetic Indexing and Search

• Audio → phonetic index
• Search term + phonetic index → hits
• No out-of-vocabulary problem
• Fast (>200 xRT)
• Robust to noisy channels and lossy codecs
• Confidence score allows selection of operating point in precision vs. recall curve

(Source: “Fast decoding for open vocabulary spoken term detection” by IBM Research, in Proceedings of NAACL/HLT 2009)
Phonetic Analysis

<table>
<thead>
<tr>
<th>Language</th>
<th>Phonemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA English</td>
<td>39</td>
</tr>
<tr>
<td>UK English</td>
<td>44</td>
</tr>
<tr>
<td>Spanish</td>
<td>29</td>
</tr>
<tr>
<td>Japanese</td>
<td>29</td>
</tr>
</tbody>
</table>
Phonetic Indexing and Search

Step 1: **Indexing**

"...transfer me to the Washington office."

**Phonetic Index** (~4 MB per hour of audio)

Step 2: **Search**

**Structured Queries** capture context via Boolean and temporal operators

Powers call categorization, quality monitoring, etc

© 2010 Nexidia Inc. All rights reserved.
Research Goals

Historical speech processing challenges…

- Low accuracy of speech-to-text in call center scenario
  - Open-ended, spontaneous conversation
  - Highly varying accents
- Poor quality telephony speech
  - Encoding/compression, background noise, clipping, on-hold music
- Computationally expensive
- Cannot search new spoken phrase without having to retrain
- English-centric R&D

…have driven our Research agenda:

- **Performance**
  - Increase speed of Index and Search
  - Minimize size of Phonetic Index
- **Accuracy**
  - Maximize the precision and recall of search results
  - Increase robustness to channel variations (e.g., 6-kHz VOX or low-bit-rate mp3)
  - Improve robustness to speaker and dialectal variations
  - Effectively handle non-speech artifacts
  - Work consistently across languages
- **Scalability for enterprise deployments**
  - Optimize algorithms to scale within a core, across cores in a server, and across servers
  - Deliver high performance in small footprint for lowest TCO
- **Technologies beyond spoken content analysis**
  - **Classifiers**: Voice Activity, Music, Gender, DTMF
  - **Language ID**: User-trainable classifier framework
  - **Clip spotting**: to identify repeated on-hold speech
  - **Transcript sync**: to robustly time-stamp transcript

© 2010 Nexidia Inc. All rights reserved.
What to look for in a Speech Analytics Solution

Accuracy

• Robustness to background noise, lossy codecs, speaker variations

Scalability

Flexible & Rich Toolbox

Language ID, voice, music, DTMF detectors

Relevant call playback & ad-hoc search
<table>
<thead>
<tr>
<th>Technology</th>
<th>Patent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core phonetic indexing and search technology (creation of PAT file, search of PAT file, sorting of hits by score)</td>
<td>“Phonetic Searching” Patent 7,263,484; issued Aug 28, 2007</td>
</tr>
<tr>
<td>Indexing and search covering both forward and backward directions in time</td>
<td>“Phonetic Searching” Patent 7,324,939; issued Jan 29, 2008</td>
</tr>
<tr>
<td>Search via linguistic search term plus phonetic search term or voice command</td>
<td>“Phonetic Searching” Patent 7,475,065; issued Jan 6, 2009</td>
</tr>
<tr>
<td>Application of phonetic indexing and search to robustly time-align a transcript with audio</td>
<td>“Transcript Alignment” Patent 7,231,351; issued Jun 12, 2007</td>
</tr>
<tr>
<td>Search of phonetic index and/or textual transcript</td>
<td>“Transcript Alignment” Patent 7,487,086; issued Feb 3, 2009</td>
</tr>
<tr>
<td>Structured query normalization: statistical modeling of score distributions of putative hits to characterize and reduce false alarms</td>
<td>“Word Spotting Score Normalization” Patent 7,650,282; issued Jan 19, 2010</td>
</tr>
</tbody>
</table>
Applications of Speech Analytics
Applications of Speech Analytics

- **Call center**
  - Business processes
    - First call resolution
    - Customer experience
    - Call categorization
  - Agent performance
    - Real-time agent assist
    - Tracking KPIs at the agent / supervisor / site levels
    - Language assessor
  - Compliance
  - Market intelligence
  - Tactical analysis

- **Legal market**
  - E-discovery
  - Compliance
  - Audio forensics

- **Rich media**
  - Video search
  - Tagging and syndication of audiovisual assets
  - Transcript synchronization
  - Audio conform analysis

© 2010 Nexidia Inc. All rights reserved.
Video Search Plug-in for Apple Final Cut Pro
“A process that used to take hours has literally been reduced to minutes and is probably one of the greatest productivity gains of any new NLE feature to come along in years.”

Oliver Peters, *Videography* magazine

© 2010 Nexidia Inc. All rights reserved.
Transcript Sync for MSNBC

Select clips from our ‘Explore’ playlist for a rich video experience

1. Easily navigate the video using the transcript
2. Find important moments using keywords
3. Clip and embed your own video edits from the transcript

© 2010 Nexidia Inc. All rights reserved.
Solutions for the Call Center

**Historical Analytics**
- Enhanced Reporting & Analytics
- Value Added Consulting

**Real-Time Monitoring**
- Speech-Driven Real-Time Event Generator
- Contact routing
- Business processes
- Self service optimization
- Agent efficiency and effectiveness
- Product / service quality
- Alerts & notifications
- Live agent assistance
- Exception management
- Proactive customer service

© 2010 Nexidia Inc. All rights reserved.
### Nexidia Products for the Call Center

<table>
<thead>
<tr>
<th>Products</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic Topic &amp; Trend</td>
<td>Identification of emerging topics</td>
<td>Enterprise Speech Intelligence (“ESI”)</td>
<td>Real Time Monitoring (“RTM”)</td>
<td>Language Assessor (“LA”)</td>
<td>Audio Finder</td>
</tr>
<tr>
<td>Discovery (“ESP”)</td>
<td></td>
<td>Root-cause analysis with historical data</td>
<td>Scanning live audio and triggering events in real time</td>
<td>Measuring speaker pronunciation and fluency</td>
<td>Stand-alone audio forensics application</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technologies</td>
<td>Language Packs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workbench SDK</td>
<td>Index</td>
<td>Search</td>
<td>Scanning</td>
<td>Language ID</td>
<td>Clip Spotting</td>
</tr>
<tr>
<td>(Linux, Windows, OSX)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Improving Business Processes in the Call Center

- Call categorization
  - Cross-industry and industry-specific call drivers
- Customer experience
  - Call transfers, escalations, repeat callers, agitated customers
- First call resolution
- Cost reduction
  - Call volume, average handle time, talk vs. non-talk time
- Sales increase
  - Cross-sell/up-sell conversion rates
Nexidia ESI

Call Driver Dashboards

Organizational Dashboards

Call Driver
Talk Time Analysis
Correlation
Trend
Analysis of Means

© 2010 Nexidia Inc. All rights reserved.
Nexidia ESI: Forensic Search
## Call Categorization Libraries

### Customer Experience
- First call resolution
- Transfers
- Negative experience
- Hold
- Call backs

### Health Care
- Claims
- Benefits
- Enrollment

### Wireless
- Plans and usage
- Troubleshooting Coverage

### Health Care
- Claims
- Benefits
- Enrollment

### Third Party Collections
- Mini-Miranda
- Collections effectiveness
- Disclosures

### Generic Call Drivers
- Billing
- Compliance
- Competitors
- Customer churn

### Utilities
- Outages
- Risk management

### Telecom
- Outages
- Dispatch
Applying Speech Analytics

A. Automatic Discovery

B. Mine Caller Intent

C. Identify Root Cause

D. Monetize

E. Develop Action Plan

F. Continuously Measure

© 2010 Nexidia Inc. All rights reserved.
“First call resolution rate is 83%”

Questions demanding insight about the remaining 17%

• What are the topics among repeat callers?
• What are the broken processes that are driving repeat calls?
• What specific agents are exhibiting behavior driving repeat calls?
• Are there calling cycles that can provide insight into behaviors?
• Who are my most frequent callers and why?
First Call Resolution Analysis

1. Data Sources

   - CRM
   - WFM
   - ACD
   - Recorder

   - Integration to customer data

2. Assemble Call Series (repeat callers)

   - Customer Call History
     - 87287118
     - 90823012
     - 20802101
     - 02107172
     - 67346688
     - 71872971

   - Flexible selection of repeat caller profiles (repeat, habitual, chronic)

3. Analyze Call Drivers

   - Powerful analysis of call drivers

© 2010 Nexidia Inc. All rights reserved.
Defining Repeat Callers

- Allows analyst to easily define repeat caller profiles:
  - Repeat callers (>2 calls in 30 days)
  - Habitual callers (>20 calls in 6 months)
  - Service failure repeat callers (>5 calls in 2 days)

What metadata is used to identify the customer?

What constitutes a repeat caller (e.g., ‘n’ calls in ‘m’ days)?

Should call series be extended by calls that fall (closely) outside the definition of a series?

© 2010 Nexidia Inc. All rights reserved.
Topics Report

- Why are customers contacting us multiple times?
- What are the call drivers and related topics?

Relationships Report

- What was the initial call reason?
- What are subsequent call reasons?
Call History Playback and Ad-hoc Search

Review and search the entire call history for that customer

<table>
<thead>
<tr>
<th>Customer Id</th>
<th>Calls</th>
<th>Total Duration</th>
<th>Start Date</th>
<th>End Date</th>
<th>Exclude this Customer Id</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-15240-25</td>
<td>12</td>
<td>4050 Seconds</td>
<td>1/1/2008 9:57:02 AM</td>
<td>1/10/2008 12:20:01 PM</td>
<td>Exclude Customer Id</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Score</th>
<th>Media Filename</th>
<th>Date/Time</th>
<th>Duration</th>
<th>Non-Tal...</th>
<th>Agent</th>
<th>Supervisor</th>
<th>Site</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CALL-28368V92</td>
<td>1/1/2008 9:57:02 AM</td>
<td>00:08:11</td>
<td>30 %</td>
<td>Greg Frankel</td>
<td>Anne Bonds</td>
<td>Atlanta</td>
</tr>
<tr>
<td></td>
<td>CALL-4685V37</td>
<td>1/4/2008 6:52:01 PM</td>
<td>00:14:28</td>
<td>30 %</td>
<td>Joe Patrick</td>
<td>Preston Williams</td>
<td>Atlanta</td>
</tr>
<tr>
<td></td>
<td>CALL-23767V7</td>
<td>1/5/2008 3:19:30 PM</td>
<td>00:01:27</td>
<td>12 %</td>
<td>Andre Osborne</td>
<td>Brian Archer</td>
<td>London</td>
</tr>
<tr>
<td></td>
<td>CALL-17701V90</td>
<td>1/6/2008 8:38:48 AM</td>
<td>00:10:52</td>
<td>62 %</td>
<td>Joe Ward</td>
<td>Terri Abrams</td>
<td>London</td>
</tr>
<tr>
<td></td>
<td>CALL-21705V23</td>
<td>1/6/2008 4:20:37 PM</td>
<td>00:07:53</td>
<td>32 %</td>
<td>Greg Frankel</td>
<td>Anne Bonds</td>
<td>Atlanta</td>
</tr>
<tr>
<td></td>
<td>CALL-5625V61</td>
<td>1/8/2008 6:46:53 PM</td>
<td>00:06:39</td>
<td>44 %</td>
<td>Joey Elias</td>
<td>Marcia White</td>
<td>Atlanta</td>
</tr>
<tr>
<td></td>
<td>CALL-3302V546</td>
<td>1/9/2008 12:08:26 PM</td>
<td>00:08:47</td>
<td>57 %</td>
<td>Andy Frankel</td>
<td>Shelley Brown</td>
<td>Atlanta</td>
</tr>
<tr>
<td></td>
<td>CALL-10406V284</td>
<td>1/9/2008 12:10:14 PM</td>
<td>00:03:56</td>
<td>7 %</td>
<td>Andy Frankel</td>
<td>Shelley Brown</td>
<td>Atlanta</td>
</tr>
<tr>
<td></td>
<td>CALL-28394V336</td>
<td>1/9/2008 2:25:59 PM</td>
<td>00:01:21</td>
<td>6 %</td>
<td>Anthony Feld...</td>
<td>Brian Archer</td>
<td>London</td>
</tr>
<tr>
<td></td>
<td>CALL-2814V434</td>
<td>1/9/2008 5:55:42 PM</td>
<td>00:00:06</td>
<td>0 %</td>
<td>Christopher Pitts</td>
<td>Anne Bonds</td>
<td>Atlanta</td>
</tr>
<tr>
<td></td>
<td>CALL-5473V220</td>
<td>1/10/2008 9:52:16 PM</td>
<td>00:01:02</td>
<td>0 %</td>
<td>Jane Carlton</td>
<td>Marcia White</td>
<td>Atlanta</td>
</tr>
<tr>
<td></td>
<td>CALL-31232V552</td>
<td>1/10/2008 12:20:01 PM</td>
<td>00:02:48</td>
<td>2 %</td>
<td>Jane Carlton</td>
<td>Marcia White</td>
<td>Atlanta</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Customer Id</th>
<th>Calls</th>
<th>Total Duration</th>
<th>Start Date</th>
<th>End Date</th>
<th>Exclude this Customer Id</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-71424-84</td>
<td>11</td>
<td>5165 Seconds</td>
<td>1/1/2008 2:07:06 PM</td>
<td>1/10/2008 3:29:52 PM</td>
<td>Exclude Customer Id</td>
</tr>
<tr>
<td>C-06745-70</td>
<td>10</td>
<td>3213 Seconds</td>
<td>1/1/2008 12:45:06 PM</td>
<td>1/10/2008 3:50:44 PM</td>
<td>Exclude Customer Id</td>
</tr>
<tr>
<td>C-09110-28</td>
<td>10</td>
<td>1515 Seconds</td>
<td>1/2/2008 2:36:45 PM</td>
<td>1/10/2008 5:09:11 PM</td>
<td>Exclude Customer Id</td>
</tr>
</tbody>
</table>
Nexidia ESI—Quality

**What is ESI—Quality?**

- Component of Nexidia’s flagship Enterprise Speech Intelligence (ESI) software suite
- Allows quality management to take on strategic role within the organization
- Measures Quality Initiatives at the site, team, and agent levels based on 100% of available calls
- Helps drive agent coaching and performance improvements in line with corporate objectives

**How does ESI—Quality work?**

- Definition of Quality Initiatives
  - “AHT for billing calls should be under 5 minutes”
  - Easy-to-use interface helps define measurable performance objectives, applicable set of calls, and expected goals
- Categorization of calls and tracking of objectives
  - ESI—Quality automatically categorizes all calls and measures each agent against the applicable Quality Initiatives
- Presentation of results in Quality Portal
  - Managers and team leads log onto Quality Portal to track performance of each agent, team, or site against expected goals
  - Relevant calls and evaluation forms are just one click away
## Performance on Quality Initiatives by Agent

The image shows a quality portal with a table listing agents and their performance metrics across various quality initiatives. The table includes columns for Resource, Dissatisfied, Transfers Excessive, Transfers ATT, Transfers Non-Talk, Connectivity ATT, Email ATT, Security ATT, and All Calls. Each agent has a status indicating whether they are dissatisfied or not for each metric.

### Quality Portal


### Drill down to calls

Agents with a dissatisfied status may require further investigation into the calls they handled to identify specific issues.
From Quality Portal to Forensic Search

Targeted Playback

Evaluation Launcher

It's not just frustration – it's outrage!

© 2010 Nexidia Inc. All rights reserved.
Evaluation Form

Configurable Evaluation Form

Refers back to specific call(s) being reviewed

Integrates with existing coaching system
Quality Initiatives are easily defined:

1. **Target Call Sets** determine applicable media
   - Powerful set definition combining query and metadata results
   - E.g., modem transfer calls

2. **Metric** captures specific objective
   - E.g., average non-talk time

3. **Acceptability Parameters** determine performance icons in Quality Portal
   - E.g., less than 3 minutes
Nexidia ESI—Capture

What is ESI—Capture?

• Real-time call capture and monitoring solution
• Allows access to audio without requiring legacy recorders
• Powers real-time applications such as Agent Assist for improved agent efficiency
• Can record calls and metadata

How does ESI—Capture work?

• Taps into IP telephony switch
  • Requires Avaya Application Enablement Services (AES) 4.1 and Communication Manager (CM) 4.0
  • Avaya DevConnect certified
  • Integration with Cisco, other platforms in progress
• Scans 100s of live streams in real time
  • Applies scanning engine to identify key phrases with subsecond latency
  • Provides event notification API
• Agent desktop alerting
  • Nexidia Agent Assist provides relevant alerts and knowledge articles
  • ESI—Capture also integrated with Convergys Dynamic Decisioning System and Quick Connect
1. Monitors conversations in real time

2. Identifies candidate knowledge articles based on trigger-action rules

3. Suggests article to agent and allows them to discard or launch
Nexidia ESP

What is Nexidia ESP?

• Automatic topic and trend discovery for proactive speech analytics
• Provides **top topics** (snapshot), **hot topics** (trends), and **topic relationships**
• Component of Nexidia’s flagship Enterprise Speech Intelligence (ESI) software suite

How does Nexidia ESP work?

• **Content discovery**
  - Crawls web sites, file system for text documents
  - Applies natural language processing techniques to extract relevant phrases
• **Media search**
  - Applies optimized algorithm to search for large set of phrases at once
• **Trend Analysis & Visualizations**
  - **Word clouds** present top phrases and hot/cold phrases
  - **Hyperbolic trees** present related phrases
  - Audio just one click away
ESP Visualizations

Top Phrases

Animated Trends

Hot/Cold Phrases of the Day/Week/Month

Related Topics

© 2010 Nexidia Inc. All rights reserved.
**Top Movers:** Colombia, Manuel Zelaya, Roberto Micheletti, Honduras, interim government; swine flu, WHO

**Related to Micheletti:** Zelaya, Honduran, curfew, political crisis

© 2010 Nexidia Inc. All rights reserved.

*Voice of America* broadcasts, May-June vs. July-August, 2009
Topic Discovery and Trending in Arabic
Topics Discovery and Trending in Arabic translated into English

Ali Abdullah Saleh Concerns Decisive military action Effects Elements of the PKK
Fourteen Index International relief organizations Islamic Dawa Party
Military operations Ramadan Recent Some media
Supervised the deployment of The month of Ramadan Two thousand and seven Unexpected
White House spokesman Yemeni army Yemeni government

Background protests Bethlehem Bethlehem Bethlehem Faten Car bomb Central Council and the Revolutionary Movement
Counter-Terrorism Fatah Fatah in the Gaza Strip For election to the Central Committee
Formation of the government Human Rights In charge of Saud al-Hariri Interview Live Manwan Barghouti
Members of the Central Committee More details More than three Negotiation Of the Central Committee
Telephone
The Central Committee and Revolutionary Council
Related Topics in Arabic

*Alhurra* broadcasts, first half of August vs. second half of August, 2009
What is Language Assessor?

- Language Assessor is a solution that applies phonetic-based speech analytics to automatically evaluate spoken language skills

- Used in the call center for
  - **Agent screening**: Pre-screen candidates based on language skills in a fast and cost-effective manner
  - **Ongoing agent evaluation**: Track progress of language skills and vary target goals according to tenure

- Key benefits
  - Provides unbiased, consistent assessment of spoken language skills
  - Reduces infrastructure needs
  - Curtails sharing and rehearsing of tests

How does Language Assessor work?

- **Script reading**
  - Applicant uses web-based interface to view and record script

- **Automated assessment**
  - Solution processes recording to compute pronunciation, fluency, and pace scores
  - **Pronunciation assessment**: Models how closely individual speaker follows standard pronunciation
  - **Fluency assessment**: Models lack of inappropriate pauses, false starts, repetitions, hesitations, and other speech disfluencies
  - **Pace assessment**: Computes speaking rate relative to applicant pool
Administrator Review Process

Core Analyzer

- Pass
- Review
- Fail
- Retest

Dashboard Reports

Assessment Details and Review

© 2010 Nexidia Inc. All rights reserved.
Sample Dashboard Reports

Applicants Summary report

- Total Applicants: 5122
- Recording Complete: 5065
- Pending Recording: 57
- Pending Review: 56
- Review Complete: 5009
- Rejected: 633
- Other: 0
- Passed: 3547
- Review Complete: 56
- Failed: 2709
- Auto: 836
- Manual: 43
- Retest: 829
- Rejected: 829
- Other: 0
- Total: 69%

Score Histogram report

- Total Applicants: 5122
- Recording Complete: 5065
- Pending Recording: 57
- Pending Review: 56
- Review Complete: 5009
- Rejected: 633
- Other: 0
- Passed: 3547
- Review Complete: 56
- Failed: 2709
- Auto: 836
- Manual: 43
- Retest: 829
- Rejected: 829
- Other: 0
- Total: 69%

Count of Applicants Recording Complete By Score

- Score (In 5 point groups)
  - 0-5: Pending Review
  - 6-10: Passed
  - 11-15: Rejected
  - 16-20: Retest
- Overview of applicants received and their status
- Processed applicants summarized by Score and current status
- Number of applicants received each day over specified time period (max 30 days)
Applicants sorted by different metrics

Phrase-by-phrase evaluation and playback of recording

Detailed Phoneme Performance chart suitable for accent coaching
Conclusion
Speech Analytics Ecosystem for the Call Center

- Real-Time Agent Assist
- Phonetic Index
- Structured Queries
- Dashboards and Reports
- Live Capture from Switch
- Extraction from Recorder
- Language Assessor
- Quality Monitoring
- Topic Discovery and Trending
- Forensic Search

© 2010 Nexidia Inc. All rights reserved.
Nexidia’s Key Differentiators

Core technology
• Accuracy
• Robustness
• Scalability
• Integrations with recording platforms
• 30+ languages
• Detectors and classifiers

Analytics methodology
• Query development
• Analytics framework

Innovative products
• ESP
• ESI
• ESI Quality / Capture / Agent Assist
• Language Assessor
• Enrich
• Syndicate
• ‘get’ Final Cut Pro plug-in
• Audio Finder