Multimodality in Mobile Banking

Matthew Yuschik, PhD.
Mobile Solutions Architect
Global Consumer Technology / CTO  R&D
Citibank, N.A.
and
AVIOS Board of Directors

Outline
What is Multimodal Mobile Banking?
High Level Needs Assessment
Multimodal User Interface Design
Biometrics
MultiFactor Authentication
Q/A
What is Multimodal Mobile Banking?

**Transparent Navigation and Data Entry => UX and Back End Integration**
Banking Transactions by speaking or touching a Smartphone, Then viewing or hearing the results of the dialog

**Biometric Authentication => Non-intrusive, transparent Security**
Voice Recognition
Facial Recognition
High-Value Services for Smartphone Users
- Entertainment
- Location based-services
- News and sports updates
- Social Networking
- Transaction-based services => banking
  - 2012 - 24% of adults used mobile banking
  - 2013 – 35% of adults used mobile banking

Combined Modes on a Smartphone
- Existing web page (GUI) or IVR apps (VUI)
- Native Features
- Fixed set of transactions

Strengths and Challenges of a MMUI
- Reference application to identify issues
- Cognitive, Behavioral, Kinesthetic constraints
- Error detection and Correction

Prof. Dan Siewiorek of CMU shares his vision of Smartphone evolution.
(IEEE Spectrum, Sept 2012, Generation Smartphone)
Multimodal Mobile Banking - Use Cases

“Account balances”

“Show me my 5 most recent transactions.”

“Pay $200 from checking to my Verizon bill next Friday”

“Where is the nearest ATM?”

Typical Use Cases:
Show Balances
Recent Transactions
Payment and Transfer
Find ATM

MMUI = GUI + VUI
Say what you want, correct when needed.

Toolkits and Development Environments
Rapid prototyping, Usability Testing
Flexible Transaction Flow – customer control of alternative paths
Compendium at http://www.avios.org/app_dev.htm

Mobile Frameworks
Browser-based languages (HTML5, JQuery, JQuery Mobile, Java Script)
Hybrid / Cloud Solution – Application on server, rendered on Smartphone client
Nuance NINA (Android and iOS) starter kit + demo applications
Native Features through Cordova (PhoneGap) or Voxeo Tropo
Multimodal Banking Transactions – Performance Analysis

“Transfer $100 from my savings to checking on Friday”

Time and Effort Comparison

<table>
<thead>
<tr>
<th>Service</th>
<th>Steps</th>
<th>Web</th>
<th>Tablet</th>
<th>MM</th>
<th>Steps</th>
<th>Tablet</th>
<th>MM</th>
<th>time (secs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Account Balance</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recent Transactions</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td>1.3</td>
<td>1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Make a Payment</td>
<td>13</td>
<td>14</td>
<td>3</td>
<td></td>
<td>23.1</td>
<td>5.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Find an ATM</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>2.1</td>
<td>1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Transfer Funds]</td>
<td>13</td>
<td>14</td>
<td>3</td>
<td></td>
<td>23.1</td>
<td>5.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fitt’s Law \( \Rightarrow \) Time for Movement (speed, accuracy trade-off):
\[ T \sim \log_2 \left( \frac{2d}{W} \right) \]
close, big targets are reached faster than small targets farther away

Conclusion:

Highest payback for voice in multi-step transactions.

Key Value Drivers:

- Ubiquity of Smartphone – Untapped Market
- Anywhere, anytime – Customer retention
- Ease of use, reduced errors – Customer Satisfaction
- Consistent UI/UX – Faster Time To Market
- Differentiation – Market capture, rewards, cards
- Shorter transaction time – Lower operation costs
- Better self service – Efficient utilization of back-end resources

Tablet

Multimodal

14 steps (if new date)
9 steps (if today’s date)

3 steps
Multimodal Mobile Design - UI/UX Mental Model

**Terminology**
Common visual layouts, verbal consistency, reduced clutter, branding
Colloquial Grammar, sentence structure, reference ("this one")

**Increased Complexity => Expectation, Cognitive Load, Learning Transfer**
Utilize modal preferences – say what you want, read the response
Follow Natural Conversation and Behavioral Flow, attention and memory
User Control – monitor what is done most often, and how; "stickiness"

**Dialog Flow**
Security - seamless
Streamlines - “jump” vs multiple steps
Fewer errors since fewer GUI navigation steps

**Error Correction**
“Breadcrumbs”, Detection, highlighting
Alternative modes of content entry (multimodal)
Authentication Biometrics – High Level Description

Fingerprint - **Image**
PC and smartphone - swipe/scan, explicit, fussy, spoofable
1-1 or 1-from-N match to userID, or identify from a large DB

**Signature / Gesture - Accelerometer**
Smartphone - cadence, [Android prelim. investigation]

**Face - Image**
Transparent - non-intrusive, no PIN or KBA
Performance authorize <1sec, EERate <1.5%
Consistent UI - any device, any channel
Easy enrollment - fast, effortless

**Voice - .wav file**
Reduced Fraud - increased security, EERate <2%, FA to .1%
Multiple Factors - initial and within, explicit or transparent
Multi-channel - IVR, web, mobile phone, VoiP
No device costs - server license fee, Web Services APIs
Language Independent - international, personal passphrase
Tuning Tools - reduced operations cost, browser-based tools
Voice Authentication – Enrollment and Verification

**Evaluation Studio**

- **Easy to Enroll**
- **Easy to Verify**
- Set Threshold as trade-off
- Extra Factor(s) if needed
Voice Authentication – Measuring User Experience

Preference Scores
A questionnaire is given after the Trial.
The 5-Point Mean Opinion Score indicates agreement with the statement,

<table>
<thead>
<tr>
<th>Statement</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>The service was easy to use.</td>
<td></td>
</tr>
<tr>
<td>The prompts were easy to follow.</td>
<td></td>
</tr>
<tr>
<td>I felt comfortable using the service.</td>
<td></td>
</tr>
<tr>
<td>I didn't need any help.</td>
<td></td>
</tr>
<tr>
<td>It was easy for me to construct my voiceprint.</td>
<td></td>
</tr>
<tr>
<td>I don't mind having a secret question for backup.</td>
<td></td>
</tr>
<tr>
<td>It was easy to choose a secret question and its answer.</td>
<td></td>
</tr>
<tr>
<td>I'd prefer more secret questions to choose from.</td>
<td></td>
</tr>
<tr>
<td>It was easy to have my voice verified.</td>
<td></td>
</tr>
<tr>
<td>I like using a voiceprint better than entering a PIN</td>
<td></td>
</tr>
</tbody>
</table>

Interpretation
- easy to use, comfortable, and users didn’t need any help
- it easy to construct and use a voiceprint.
- there is no issue with a Secret Question, but do not have more than 5 alternative questions
- a voiceprint is slightly preferred over a PIN.
VA & Geolocation - Multi Factor Authentication

Route

Call Flow

Confidence Score

Value Drivers

Security
• MFA - fusion of 2 Factors,
• Specify Decision Rules - Site, Country, Business, etc.

Customer Focus
• Minimal UI/UX complexity (once, at initial step)
• Transparent geo (no user interaction)

IT Efficiency
• Hybrid with Native features of smartphone
• Geo-centers and geo-fence distances are tunable
• No delay for geolocation information
• No 3rd party components, integration, tuning, support

1900 miles - Some network and road effects
11 mobile sites: lower scores, lower SNR, more variance & rejections
1 fixed site - higher scores, no rejection
Multi-Factor Authentication - Multiple Voice Stages

Each stage is an independent test, and effects treated as independent events. More tests reduce False Acceptance, increasing security. Order of modules must be seamless.

Other Authentication Issues:
- MFA fusion rules (fuzzy logic, linear wts.)
- Other factors: Credit Card PIN, Geolocation
- Other VA processes: Directed Digits (like, OTP) with ASR for digits and VA for speaker

Security Levels – Voice compared to Data

<table>
<thead>
<tr>
<th>Risk Level</th>
<th>permissions</th>
<th>Data</th>
<th>Voice</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Basic Info: acct balance, order checks, stop payment, pay bill</td>
<td>Password, PIN</td>
<td>PIN &amp; voiceprint</td>
</tr>
<tr>
<td>2</td>
<td>new payee, withdrawal, transfer change PW, change address</td>
<td>Knowledge-Based (secret Q/A)</td>
<td>Liveness Detection Device or geolocation</td>
</tr>
<tr>
<td>3</td>
<td>High risk transfer, trades, loan payments</td>
<td>OTP, SecureCard</td>
<td>softwareOTP = Random Numbers</td>
</tr>
</tbody>
</table>
Multimodal Mobile Banking and Mobile Wallet prototypes are being tested to identify strengths, mitigate confusions, and seamlessly direct the user back on the success path of Digital Banking.

Questions??

Thank you
Dr. Matthew Yuschik designs, develops and tests Multimodal Mobile Banking UX/IU for Citibank’s R&D group. This involves voice navigation, voice authentication, transaction analysis, and usability. He field-trialed a Multimodal Call Center Workstation in the US, and developed voice-activated voice mail for 7 European wireless providers. Matt has 9 patents, with 3 pending. He has been on the AVIOS Board for 8 years, and is currently the Treasurer.

matthew.yuschik@citi.com