Call Browser: A System to Improve the Caller Experience by Analyzing Live Calls End-to-End

Bernhard Suhm
AVOKE Caller Experience Analytics
BBN Technologies
10 Moulton Street, Cambridge MA 02138
bsuhm@bbn.com – Tel.: 617-873-4875

Pat Peterson
AVOKE Caller Experience Analytics
BBN Technologies
10 Moulton Street, Cambridge MA 02138
patp@bbn.com – Tel.: 617-873-2670

ABSTRACT
This paper describes a system that empowers practitioners to substantially improve the user experience with call center automation and agents. Unlike other approaches we analyze the caller experience in live calls end-to-end, from dialing to hangup. A web-based solution, the Call Browser provides access to hundreds or thousands of live end-to-end calls, and empowers usability practitioners and call-center analysts to systematically and efficiently evaluate the caller experience and identify usability issues. Case studies from our consulting practice illustrate how this approach reveals issues that remain hidden to traditional methods, such as log analyses, lab user studies, focus groups, and design guidelines.

Author Keywords
Interactive Voice Response (IVR), Voice User Interfaces (VUI), user experience, call center analytics, data-driven usability optimization, speech recognition.

ACM Classification Keywords
H.5.2 [Information interfaces and presentation (e.g., HCI)]: User interfaces – user-centered design, evaluation, voice I/O; H.3.5 [Information storage and retrieval]: Online information services – web-based services; D.2.2 [Software engineering]: Design Tools and Techniques – state diagrams, user interfaces

INTRODUCTION
Everybody knows and dreads them: automated telephone systems that seem more apt to confuse and frustrate callers than to assist them in getting their questions answered. The introduction of speech technologies to the call center industry a decade ago did not turn call center automation around for the better; instead of struggling to find ways out of touch-tone menus, callers now struggle with speech recognition errors and lengthy and poorly worded messages. The GetHuman initiative (gethuman.com) has provided an outlet for caller frustration by enumerating customer service numbers, assigning a rating, and describing how to circumvent automation and reach a live agent. Why has the state-of-the art in call center automation still not improved?

Decision makers in call centers often want to help callers but lack adequate information. Standard reports can mislead them to believe that existing call center automation (referred to as "IVR" for the remainder of this paper) is performing adequately. Even if they realize that something is wrong - for example, from negative feedback in caller surveys – sufficient detail to identify the specific usability problems is often missing. Analysis of IVR logs from live calls can reveal certain usability problems, but some of the most significant IVR usability issues require an understanding of caller intent, which log analysis cannot provide. Examples include whether the menus successfully guide callers to the right place, or whether an inquiry that transferred to an agent was routine and could have been self-served more efficiently.

We have presented a methodology for optimizing call center IVRs by analyzing live calls end-to-end in previous work [1]. By analyzing the complete end-to-end caller interaction, call center analysts and usability practitioners are empowered to address difficult IVR usability problems. Moreover, by pulling together information from typically disparate caller transactions with one or more IVRs or agents across a complete call, business process and agent work flow issues can be revealed that impact the customer relationship far beyond a single call. This paper presents a system that makes it easy and efficient to apply the methodology, and empowers analysts and usability practitioners to measure and improve the complete caller experience. Unlike other approaches, the system captures the caller experience from dialing to hangup, for thousands of calls into live call center applications. Analysts and usability practitioners access and analyze end-to-end call data using a web application, the Call Browser. Usability and caller experience issues can be identified and quantified in ways that are impossible with other methods commonly used in the call center practice. We demonstrate the power

CHI 2009, April 3–9, 2009, Boston, MA, USA.

of this approach using case studies that show how we discovered and addressed caller experience and usability issues in various deployed call center applications.

RELATED WORK

Common Approaches to Call Center Usability

Several approaches to improving usability of call center automation are established in the field: design guidelines, expert audits, usability studies, recognition tuning, and focus groups.

Usability know-how for menu systems - within the call center industry frequently called IVR systems - ranges from style guides for touch-tone IVRs [2] to comprehensive collections of design best practices [3]. Industry experts apply this knowledge using techniques similar to usability walkthroughs [4]. Design guidelines are useful for generating initial designs that are good starting points for the usability cycles that follow. But guidelines tend to generate several different, plausible designs, leaving the practitioner with the challenge of determining which design works best. Best practices help practitioners avoid the mistakes that others made. But guideline-based audits reveal only obvious usability problems that may affect users. They do not provide any quantitative data, such as how many live callers are actually impacted by an issue, or how many navigate menus as the designer intended.

Lab usability studies are commonly applied to IVRs, although only a few such studies have been published [5-9]. Unlike the aforementioned audits, these studies yield some amount of quantitative data. Additionally, motivations for specific user behaviors can be explored in post experimental interviews. However, the limitations of this method forces practitioners to study only a few "high value" tasks. Further, no live call traffic is analyzed in such lab studies, and due to the typically small sample size only "big" problems are likely to be identified.

As a special case of log analyses of live call traffic mentioned earlier, cycles of "recognition tuning" are frequently employed to improve the usability of IVRs that employ speech recognition. Such tuning is aimed at improving recognition accuracy by manipulating the recognition engine, for example, by optimizing certain parameters and grammars. Grammars try to capture how users respond, but responses depend on how they are prompted. Recognition tuning analyzes responses to specific prompts without the dialog context, thus missing issues where recognition and prompt sequence interact. Recognition tuning, and log analysis in general, therefore are necessary, but not sufficient, means for improving usability of call center automation.

Call centers frequently employ focus groups to identify issues with existing systems, and to generate ideas for new applications of call center automation. As with usability studies, focus groups allow the mediator to explore caller thinking and motivations, but focus groups yield little quantitative data.

We propose that a data-driven, quantitative approach can substantially improve the practitioner's ability to address poor usability of many deployed call center IVRs. Key to this approach is the way we capture caller experience data, discussed in the next two subsections.

Common Methods for Capturing the Caller Experience

Discussing the various kinds of caller experience data available to call centers requires an understanding of the various segments that calls comprise: IVR, queue, and agent segments. These segments are illustrated in Figure 1 (above the waveform), along with various common methods for gathering data about the caller experience.

Reviewing the caller experience from dialing to hangup, the caller first interacts with IVR menus, data entry prompts, and informational messages. Some calls end in the IVR, either after obtaining some useful information (called "self-serve"), or without it ("abandon"). Many calls leave the IVR for a live agent. Provided the caller does not abandon while waiting in queue, an agent will answer the phone and engage the caller in a dialog. Many calls end upon completion of this dialog. However, in some cases the first agent is unable to answer the caller's question and transfers the caller back into queue to wait for another agent. Some calls even have multiple transfers (see the example call in Figure 1).

Many call centers track performance metrics from each of these call segments in the live call traffic, but the data are fragmented and difficult to aggregate across multiple segments. For the IVR segment, the system often documents the event sequence for the IVR-caller interaction, which prompts were played, and how the caller responded. Log analysis tools [10] can present this information in flow diagrams or other forms for analysts to interpret. IVRs that employ speech recognition keep temporary copies of every caller utterance, along with the recognition result. Recognition tuning, which was described earlier, analyzes the utterance and result records to improve the performance of the recognition engine.

Once a call transfers out of the IVR, a telecom "switch" frequently takes back control of the call, assigning it to the proper queue for a live agent. Based on logs created by the switch, most call centers monitor how many calls transfer to a queue, the average wait time in queue, and how many calls abandon while waiting for an agent. Once an agent comes on the line, select pieces of data from the agent-caller dialog can be written to logs by the agent desktop system. In addition, some call centers capture the agent-caller dialog as an audio and/or screen recording using commercial recording solutions. However, neither desktop systems nor agent-side recording solutions capture the caller experience before or after speaking to a particular agent, such as in the IVR or in a conversation

with another agent. Even if the dialog with the second agent were captured with the same recording solution, it is difficult to match up separate recordings of two agent-caller dialogs. Further, calls can be transferred to another call center or IVR, making it even more difficult for the call center to keep track of the end-to-end caller experience.

Finally, companies sometimes use surveys to gather satisfaction feedback from their customers. The survey is typically administered either at the end of the call or in a separate callback following the initial call. Surveys are quite limited in their usefulness for quantifying the caller experience, and for conducting usability analyses, for the following two reasons: First, they represent a biased sample, because those callers who respond to the survey have extra time or are particularly motivated, for example, due to a particularly bad experience. Second, it is not possible to go back to the actual interaction that led to the bad experience, because it wasn't captured along with the survey. Hence, it is difficult to investigate the root cause for negative feedback, much less identify possible remedies.

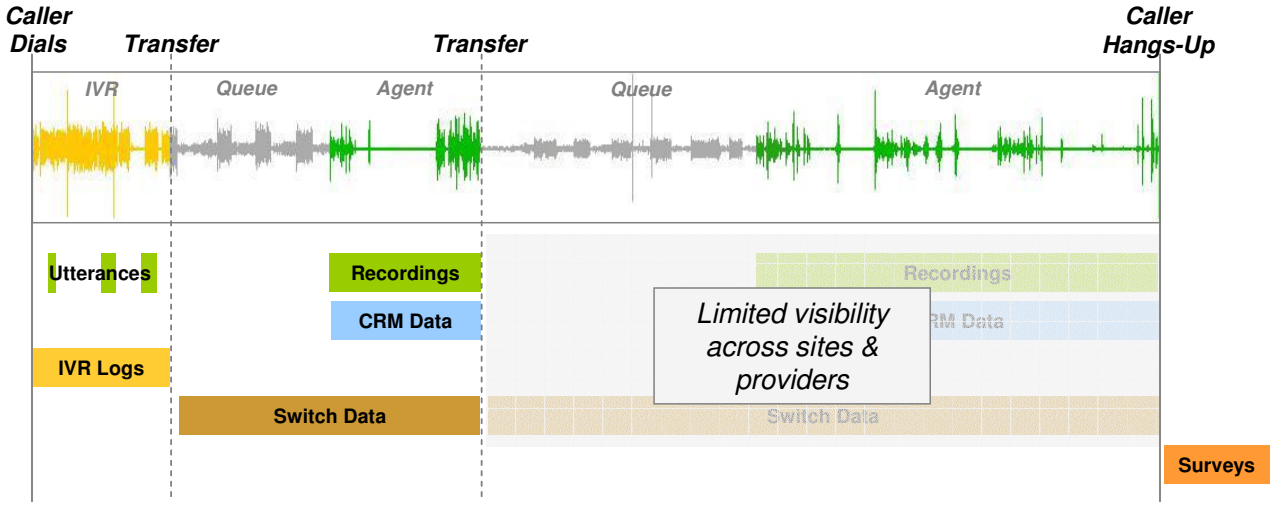

Figure 1. Data sources available for the various segments of a call (IVR, queue, agent-caller dialog).

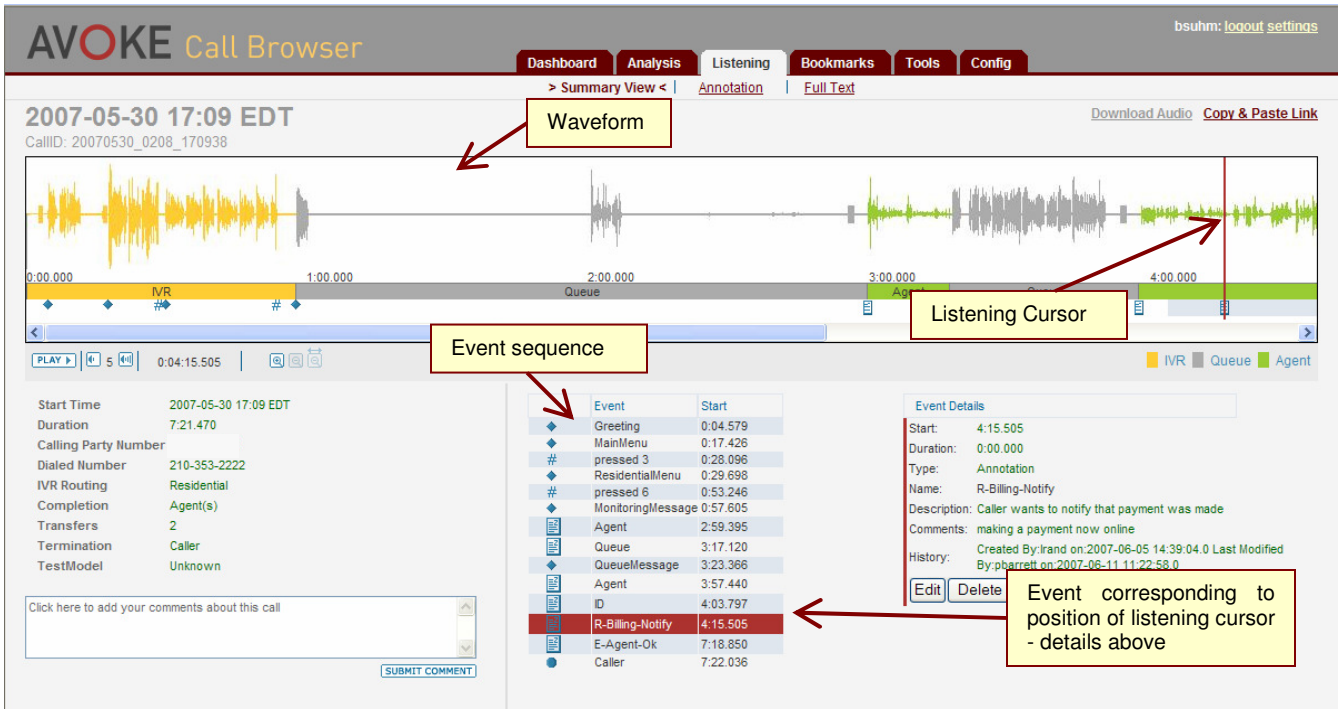

Figure 2. Call Browser listening screen showing navigational tabs on top, followed by the waveform representing the call recording, and various meta-data below on the left, such as the call occurred on 5/30/2007 at 5:09 PM and lasted 7 min 21 s. The event sequence, shown in the middle below the waveform, covers the caller's interaction with the IVR and agent-caller dialogs.

A unifying theme of the above discussion is that the fragmented nature of data available to call centers makes it difficult to identify and diagnose problems. For example, while a routine inquiry that was handled by an agent may indicate a missed opportunity to resolve a call in self-service, knowledge of the caller's behavior in the menu system is necessary to determine why the caller did not self-serve: menu navigation or self-service usability issues. The caller may not have reached the appropriate self-serve application, indicating an issue with menu navigation, or the caller may have had difficulties in the self-service application itself. Hence, the ability to analyze both the IVR event sequences and the agent-caller dialogs that follow is crucial to identifying improvements to existing self-service. - Similarly, while switch data may indicate that a call was transferred from another agent, neither switch data nor the agent-caller dialog that follows indicate whether the call was transferred because the first agent lacked training or skill, the caller picked the wrong option in the menu system, or an error in the routing table sent the call to the wrong agent pool.

Thus, a comprehensive methodology for analyzing and optimizing the caller experience should be based on end-to-end call records. The next section presents our Call Browser system for capturing and analyzing calls end-to-end.

A SYSTEM TO ANALYZE CALLER EXPERIENCES BASED ON END-TO-END RECORDINGS OF LIVE CALLS

Capturing Live Calls End-to-End

The Call Browser is an integrated network recording and analytics solution. Recording in the telephone network eliminates the need for on-premise hardware or software and enables capture of the entire caller experience – from the moment callers dial the last digit of the contact number until the moment they hang-up. The system captures a small percentage of the call center's inbound traffic. To ensure statistical validity, capturing a uniform sample is crucial. This can be achieved by having the telephone carrier divert a small percentage of the inbound traffic before it is handled by any call center equipment – another benefit of recording in the network. If the recording is performed by diverting call traffic handled on specific equipment, including the subset of traffic handled at a particular site, the sample may be biased, as certain issues (such as problems with transfers) are sometimes caused by specific equipment or its configuration.

Recording of the audio begins before the call is even connected to the originally dialed number. The caller does not notice this process – the system does not add noticeable delay or audible artifacts. The resulting recording includes the caller's interaction with all IVRs, queues and agents and continues until the caller hangs up – across any transfers that may happen in the course of a call.

Once captured, the call recording is enhanced with additional data using the following four methods. First, audio processing algorithms detect network messages, switch menus, IVR prompts, and touch-tones. Second, analytic models implemented as finite-state automata automatically classify calls along various dimensions, including which path the caller took in the IVR, at which prompt (or dialog module) the call left the IVR, number of transfers, time spent in each call segment, and how the call ended (in the IVR, in queue, or with an agent). Third, speech-to-text software transforms the entire call into searchable text. Finally, human annotators can listen to selected call recordings and insert manual annotations to characterize the caller's interaction with the IVR or analyze agent dialogs in detail. For example, the caller's reason for placing the call can be added manually by listening to the beginning of the caller-agent dialog, which is sometimes referred to as "call reason coding".

Conceptually, the call recording is stored along with automated event detection and manual annotation in a multi-dimensional database. The Call Browser provides a web interface for querying the database and accessing the data. - Next we describe how analysts and practitioners can further analyze this data.

Web-based Call Analysis

The Call Browser allows analysts and practitioners to access end-to-end call data using a secure web application. Security is important because calls contain sensitive data, such as social security numbers, credit card numbers, or personal information. Typical uses include:

1. Selecting calls of interest using various call properties ("filters"), and looking for patterns in the data by displaying it in graphs and charts
2. Listening to select calls to investigate patterns and identify root causes of usability problems, and adding manual annotations and comments to document insights in one place,
3. Sharing insights with collaborators and stakeholders

The Call Browser's Analysis and Listening screens enable these typical uses, and are described in more detail below.

Analysis Screen

The Analysis screen, shown in Figure 3 below, allows analysts to find calls that evidence problems, and to visualize call data. Since each call recording is automatically tagged with a set of properties, the analyst can hone in on a subset of calls using the filters presented in the "Call Selector". This is very powerful because the analyst is spared from hunting for examples among a large set of random calls. Instead they can inspect calls in a focused fashion, and listen to the audio by clicking on the call hyperlinks in the "Call List".

Calls of interest can also be identified by performing ad hoc word or phrase searches on the full speech-to-text transcripts of the entire call. To support systematic, statistically valid analyses, subsets of the recorded calls can be selected randomly. Finally, analysts can use the "Chart Controls" to display select data using pie, bar, and column charts. The case study in the next section illustrates the system's charting capabilities.

Listening Screen

The Listening screen allows analysts to listen to an individual call recording and to further analyze the call. An example is shown in Figure 2 above. The synchronized waveform and call event list below the waveform make it easy to navigate the audio, helping to minimize the amount of time users spend listening. Color coding indicates IVR, queue and agent segments of a call, and listeners can jump to any point in the audio just by clicking on the waveform or an event. Key data summarizing the call is displayed on the left hand side below the waveform, including when the call began, duration, which path the caller navigated in the IVR ("IVR Routing"), how the call ended ("Completion").

Manual annotations can be performed in the "Annotation" variant of the Listening Screen that displays available codes as a palette of buttons, and that allows analysts to add freeform textual comments. Manual annotations are made at specific times within a call, and thus refer to specific events in the caller experience. The analyst can assign durations to annotations, and infer average time spent on certain events within a call, which can be used to relate the impact of problems to the total time spent by the caller or the agent.

Calls can be shared with others by clicking the "Copy & Paste" hyperlink, which generates a call link that can be pasted into an email and takes any user to the same call at the exact same time within the call. The full transcript that was generated by a large vocabulary speech recognizer is accessible on the "Full Text" tab, which provides an alternative view to calls.

Supporting Collaboration

One key aspect of the Call Browser is its ability to foster collaboration among analysts and other key personnel in the call center. The interaction of automated systems, queues, and agent handling of calls is quite complex in its impact on the caller experience. Determining the root cause of problems therefore frequently requires collaboration among several subject matter experts in the call center. The following features enable collaboration:

- Adding manual annotations
- Sharing calls and charts by exchanging hyperlinks
- Adding textual comments to specific calls
- Exporting call data

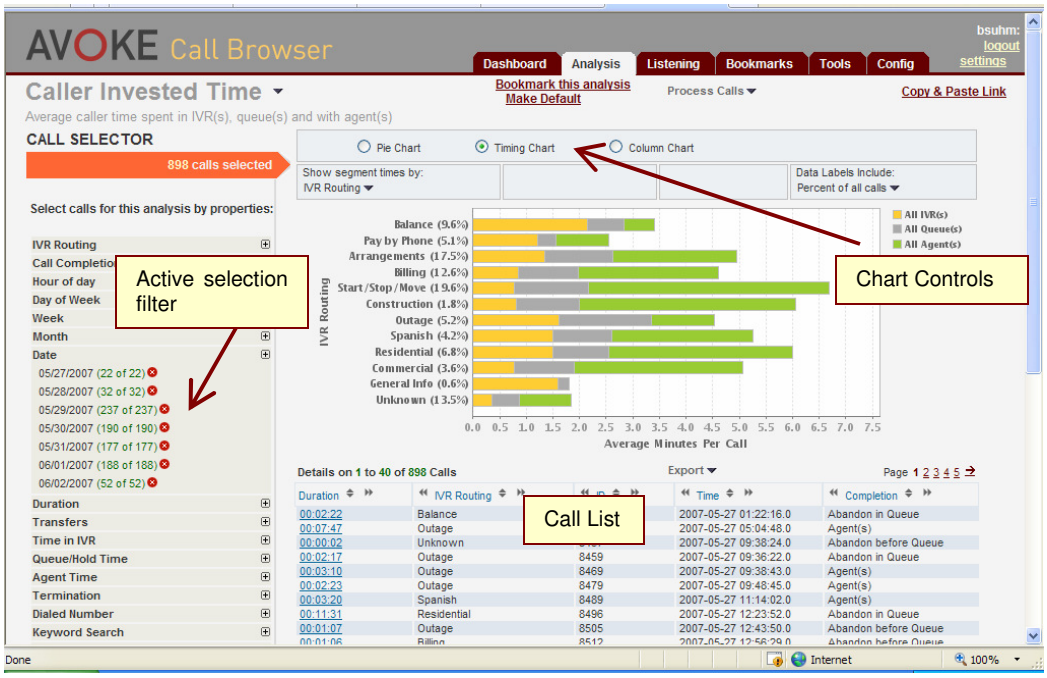

Figure 3. Analysis screen to select a subset of calls of interest using the filters available in the Selector panel on the left, to create charts, and to drill down to individual calls in the Call List below the chart.

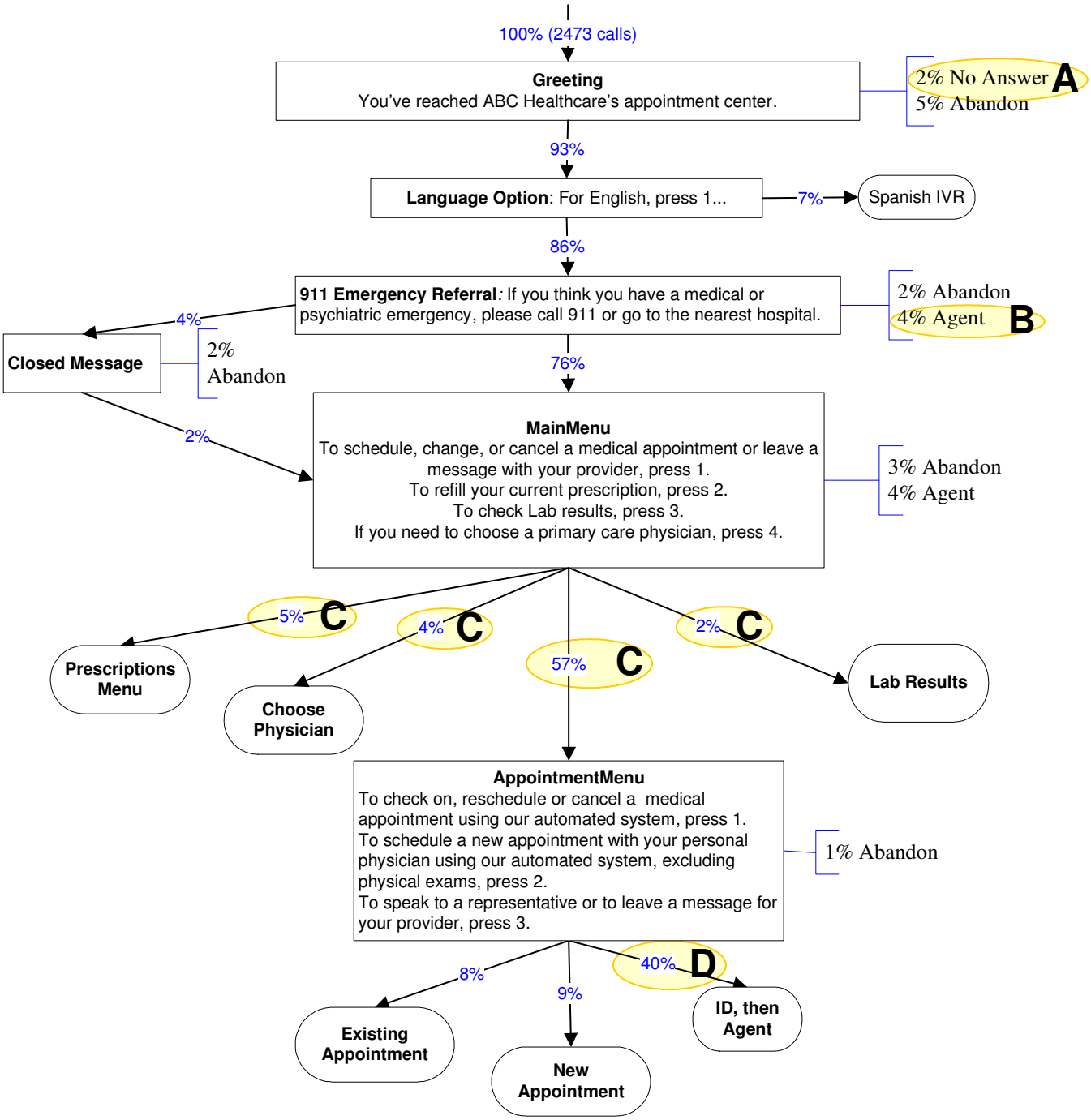

Figure 4. Caller-path diagram for Healthcare case study

UNCOVERING CALLER EXPERIENCE AND USABILITY ISSUES

This section demonstrates how the Call Browser enables systematic analysis of the end-to-end caller experience, revealing usability issues with call centers that other approaches to call center usability are unable to find. A case study of a health care and insurance provider illustrates the approach and how the system can be used to:

- Identify usability issues with the menu system
- Reveal process and performance issues that frustrate callers and waste agent time

Caller Path Analysis

Call centers frequently track caller behavior in the menu system by counting how many callers pick each option in the system, how many calls obtain some information, and whether the call ended in the IVR or transferred to an agent. While this data sheds some light on the caller behavior in the IVR, it is not sufficient to identify and diagnose usability issues. Instead, we employ a visual representation of caller behavior in the IVR.

Caller behavior in the automated menu system can be visualized effectively using a state transition diagram, which we call "Call Path" analysis. Nodes in this diagram correspond to one or more prompts of the menu system or to events in subsequent call handling, such as transfers or agent interactions. Each node and leaf is marked with the percentage of all calls that reached the node or leaf. – Details of performing caller path analyses have been described elsewhere [1] and are not unique to our system.

Using the analysis screen to plot the "IVR Exit Point" over "Call Completion" yields the raw data required for drawing a caller-path diagram. 4 above shows the caller path diagram for the Healthcare provider in the case study, representing 2473 calls that were analyzed. We identify usability problems by inspecting caller-path diagrams for areas that receive little or no caller traffic or that have high rates of abandoned calls or transfers to an agent. Some areas of interest are highlighted in Figure 4:

A. While 5% abandons at the greeting are within the normal range (a certain percentage of callers dialed the wrong number), 2% of all inbound calls were not even answered. This issue will be discussed in more detail later.

B. 4% of calls transfer to an agent at one of three upfront prompts that precede the main menu. These calls didn't even indicate their reason for calling, which typically happens with the main menu. Opportunities to direct these callers, if appropriate, to the prescription refill or appointment self-service were missed.

C. The majority (57%) of callers select the appointment option at the main menu, while all other main menu options are selected by fewer than 5% of calls. This raises questions about the menu structure: if most calls are appointment-related, why not offer appointment options at the main menu, and defer most of the little-used options that constitute the current main menu to a second menu layer?

D. Of the calls that reach the appointment menu, about 2 out of 3 (40% of total calls) select option #3 "To speak to a representative or to leave a message for your provider". Whenever the menus mention how to reach a live agent, a majority of callers will pick that option. While difficulty in reaching an agent is the most frequent complaint about menu systems, and advertising agent access therefore improves caller satisfaction, doing so prematurely leads to low usage of self-service. The analysis of menu navigation, which will be discussed in the next subsection, allows the analyst to determine whether such prompt wording indeed causes missed opportunities to self-serve callers.

Analysis of Menu Navigation

While the caller-path diagram indicates how often callers select the various options in the menu system, it does not reveal how many callers should be selecting each option, because callers may be selecting the wrong option in the menus. To determine whether callers navigate the menus correctly and pick the right option, we need to know the true reason for calling, which for most calls does not become apparent until the caller speaks with a live agent.

Annotating the reason for calls in randomly selected agent-caller dialogs allows analysts to estimate their frequency distribution. We currently employ human transcribers to annotate the call reason in the Call Browser. The distribution of the reasons for calls is a first and crucial step towards understanding why customers are calling, but it is frequently not available to designers and implementers of menu systems. In our case study, based on manual annotations of reasons in those 236 calls that were handled by a live agent, among a random call sample of 304 calls, the top three call reasons were:

- 43% (of agent-handled calls) scheduling an appointment
- 21% Leaving a message (with their provider)
- 9% Rescheduling an appointment

With call reasons in hand we can compare menu selections by callers with their true reasons for calling, and thus measure successful menu navigation as the percentage of callers selecting the correct option. Figure 5 shows as columns which option callers selected (IVR

Routing), and breaks down these columns by the reason for their call. 100% in each column represent all calls that selected a certain option, which vary greatly as shown in the caller-path chart in Figure 4. The chart empowers an analyst to identify visually which menu options are working as the columns that consist of just the matching reasons, whereas columns with many colors indicate menu options that are frequently selected incorrectly. As Figure 5 illustrates, the following options are selected correctly most of the time:

- Prescription, corresponding to option 2 in the main menu (refer back to the caller-path diagram which shows the options for the various menus)
- Existing appointment, corresponding to option 1 in the appointment menu. Cancel, check, and reschedule appointment are all matching call reasons
- Lab Results, corresponding to option 3 in the main menu
- Schedule a new appointment, corresponding to option 2 in the appointment menu

On the other hand, it becomes apparent that several options are often selected incorrectly by callers, as reflected in the more varied colors of the corresponding columns:

- Caller who leave at or before the main menu
- Leave message (option 3 in the appointment menu), which is chosen by many callers who really want to schedule a new appointment, and less for various other reasons
- Choose physician (option 4 in the main menu)

Knowing which menu options have usability problems, we infer specific solutions for improving menu navigation by analyzing the menu wording. In retrospect, it seems obvious that many callers choose the "Leave message" option because that menu choice includes the offer of speaking to a representative. It is less obvious why so many callers select the "choose physician" option incorrectly. Some wanted to schedule an appointment but knew they still needed to select their primary care physician first. However, since only 4% of all calls use this option, it doesn't have a big impact on the overall accuracy of menu navigation.

To remedy these issues we suggested exposing the frequently requested, self-servable appointment-related options at the main menu, to eliminate the second menu layer for appointments, and to consolidate all options that require agent assistance (including leaving a message for a physician) in one option of the new "other" menu", resulting in the following streamlined main menu;

For new appointments, press 1.

Existing appointments, press 2.

Lab results, 3.

More choices, 4.

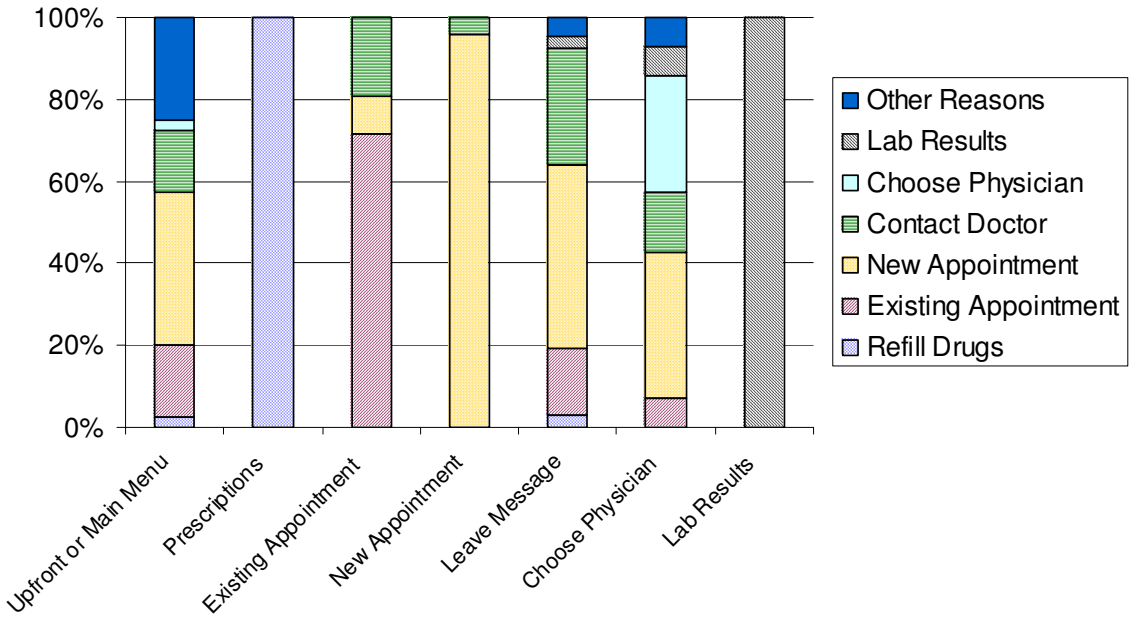

Figure 5. Analysis menu navigation accuracy. Columns represent menu selection, broken down by call reason.

Problems in the Telephone Network

Poor caller experiences can result not only from poorly designed menus and automated messages, or from how an agent handles the call, but also from problems in the telephone network. After dialing, a call may be unanswered, encounter a busy, or be dropped – not by the caller, but in the telephone network itself. Dropped calls are particularly wasteful and annoying if they occur while the caller is waiting in queue, or mid-sentence while talking with an agent. Without the capability to record calls cradle to grave, such problems are invisible to call centers. The following illustrates how automatic call classification combined with charting functionality can identify and diagnose such network issues.

Network issues can be identified based on how a call ended: who hung up first (caller or call center) and whether any signals from the network indicate an abnormal termination (call not answered, line busy, etc.). Dropped calls often terminate normally from the network's point of view – but they have in common that the call center, not the caller, hung up first. In this case study we noticed that a significant number of calls weren't answered by the call center, as pointed out earlier as issue A in the caller-path diagram.

To diagnose the issue further, we generated charts that revealed patterns in when the "no answers" were happening. Plotting the occurrences by date indicated that "no answers" were much more frequent on certain days.

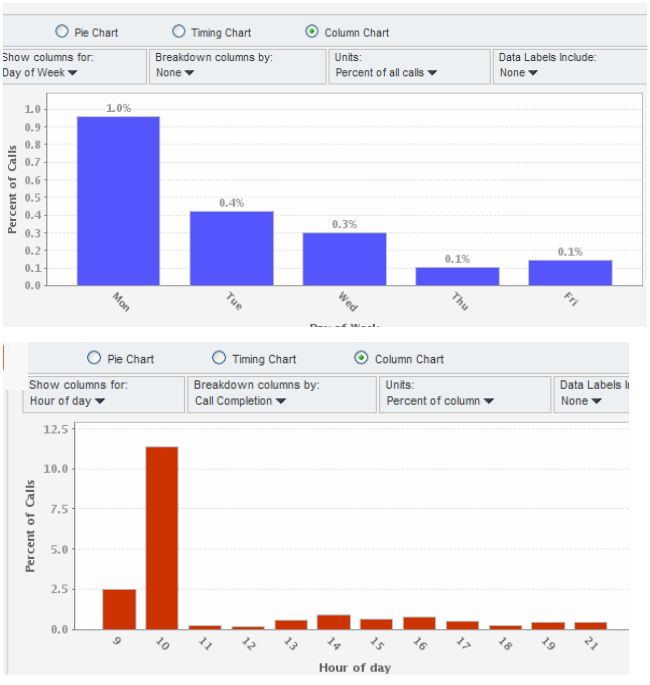

Figure 6. Occurrence of unanswered calls by day of week (upper chart) and time of day (lower chart).

By manipulating the dimension used in plotting the columns, we generated the charts in Figure 6 showing that most "no answers" occur on Mondays, and between 10-11 am EST. In discussing these patterns with the client we found out that any remaining medical appointments for the same day were released at 10 AM. Members knew that they had to call in at that time if they wanted to get an appointment that day. These business policies thus caused a spike in call traffic at 10 am, exceeding the center's capacity to answer calls. And since only emergency care is available on weekends, the call spike was particularly severe on Mondays.

Linking Survey Data with actual Caller Experiences

Mining surveys administered in call centers for usability and caller experience issues is difficult at best, because caller ratings reflect their total experience with a company's customer service, including experiences at stores, with the specific product, with technical support – not just the experience in a call center IVR that they recently called. Therefore it is impossible to establish causal relationships between their feedback and specific issues in the caller experiences. Our system offers the possibility to capture the complete interaction of the caller on a permanent record along with the survey data so that specific causal relationships can be identified. Post-call survey data can be imported, or in-call surveys can be captured in the end-to-end call recording.

DISCUSSION AND CONCLUSIONS

This paper presents a system that overcomes the limitations of traditional methods for gathering usability data from calls, and empowers practitioners to analyze the complete, end-to-end caller experience using a quantitative approach. We provide a way to let the data, instead of the "Highest Paid Person's Opinion" [11], guide us on matters impacting the caller experience. In addition to thorough checkups at particular points in time, we can perform controlled experiments in the call center environment, including A/B comparisons ([9] describes an example), and monitor systems ramping up from pilot to deployment. Our approach is strikingly similar to one recently recommended for web site usability [12].

Aside from advocating a data-driven approach, we want the HCI community to realize that standard (lab) usability testing has serious limitations in the call center context. For example, it is very difficult to simulate lack of motivation and distractions that impact live calls (such as dealing with a crying baby while making a call), and to make valid inferences about speech-enabled IVRs given the high variability in speech recognition performance.

As with all quantitative work, the Call Browser does not relieve the practitioner from paying attention to sampling, and experimental design in general, to ensure inferences are statistically valid. Aside from setting up the call

recording such that it represents a uniform sample of the total inbound traffic, any issues or recommendations derived from the sample must be interpreted with the proper statistical qualifications. We typically analyze a few thousands calls using automated processing to achieve error margins around 1%, and a few hundred calls leveraging manual call annotation to achieve error margins around 5%, which in our opinion represents a good trade-off between effort required and statistical significance. The Call Browser's automated call classification, selection, and charting techniques make it easy to analyze thousands of calls. The effort necessary for manual call annotation varies greatly depending on coding scope, domain complexity, and average call length. While simple call reason coding may average just a few minutes per call, comprehensive annotation of the end-to-end experience could take up to an hour or more. In complex domains like technical support we have come across calls that lasted several hours!

We demonstrated the effectiveness of this approach using some examples from our consulting practice. As an additional example, even though the IT department of one client at first denied any problems, the Call Browser proved that network and other issues caused them to disconnect more than 15% of their calls, which was one of the main reasons for low caller satisfaction scores. Using diagnostic insights from the Call Browser, they have since been able to reduce the disconnects to the low single percent points. Based on our recommendations, another client was able to increase the success rate of identifying callers, as well as reduce transfers, by 50%. Using the system described in this paper, we often identify specific improvement opportunities that correspond to 10% or more in operational cost savings, which almost always corresponds to significant improvements in the caller experience because reducing the time agents need to handle calls also reduces the time callers spend to get their problem or question resolved.

The Call Browser can be set up easily for any call center and it does not introduce additional legal requirements beyond what call centers have to consider when employing any other call monitoring or recording technique. Playing the commonly used monitoring notice ("Your call may be monitored or recorded for quality purposes") at the beginning of a call will satisfy these requirements in most instances.

ACKNOWLEDGMENTS

We thank Joe Alwan who created Figure 1 and helped edit the paper. We also gratefully acknowledge the feedback provided by Jack Carroll. The system described in this paper is available commercially from the AVOKE Caller Experience Analytics division of BBN Technologies (www.bbn.com/avoke).

REFERENCES

1. Suhm, B., and Peterson, P. A Data-driven Methodology for Evaluating and Optimizing Call Center IVRs. In *International Journal of Speech Technology*, 2002. Kluwer Academic Publishers, (Netherlands), Vol. 5, Issue 1, 23-38.
2. Halstead-Nussloch, R. The design of phone-based interfaces for consumers. In *International Conference for Human Factors in Computing Systems (CHI) 1989*, ACM, Vol. I, 347-352.
3. Balentine, B., and Morgan, D. P. How to build a speech recognition application. Enterprise Integration Group (San Ramon, CA), 2001.
4. Nielsen, J. Usability engineering. AP Professional (Morristown, NJ), 1993.
5. Delogu, C., Di Carlo, A., Rotundi, P., and Satori, D. A comparison between DTMF and ASR IVR services through objective and subjective evaluation. In *Interactive Voice Technology for Telecommunications Applications (IVTTA),* Italy. IEEE (1998), 145-150.
6. Niculescu, A., White, G. M., and Lan, S. S. Impact of English Regional Accents on User Acceptance of Voice User Interfaces. In *Proceedings of NordiCHI 2008*, 523 - 526.
7. Perugini, S., Anderson, T. J., and Moroney, W. F. A Study of Out-of-turn Interaction in Menu-based, IVR, Voicemail Systems. Dialog in the RAILTEL telephone-based system. In *International Conference on Human Factors in Computing Systems (CHI) 2007*, ACM, Vol. 1, 961-970.
8. Walker, M. A., Litman, D., Kamm, C., Abella, A. PARADISE: A framework for evaluating spoken dialogue agents. In *35th Annual Meeting of the Association of Computational Linguistics,* Morgan Kaufmann, 1997, 271-280.
9. Suhm, B., Bers, J., McCarthy, D., Freeman, B., Getty, D., Godfrey, K., and Peterson, P. A Comparative Study of Speech in the Call Center: Natural Language Call Routing vs. Touch-Tone Menus. In *International Conference on Human Factors in Computing Systems (CHI) 2002*, ACM, Vol. 1, 283-290.
10. Goss, K., and Gilbert, J. A Multiple Approach is Best. http://www.speechtechmag.com/Articles/Column/Talking-Tech/A-Multiple-Approach-Is-Best-36885.aspx, July, 2007.
11. Kohavi, R., Henne, R. M., Sommerfield, D. Practical Guide to controlled Experiments on the Web: Listen to Your Customers not to the HiPPO, In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2007, San Jose, CA, 959-967.
12. Kaushik, A. Web Analytics: An Hour a Day. John Wiley and Sons, 2007, 443 pages.