

Executive Summary

Natural language understanding technology holds the promise of radically innovating computer-human interfaces for disabled communities, as well as for a broad range of personal and business application users. Realizing this promise requires careful and frequent attention to how real-world users interact with this new technology as it evolves in the context of practical applications. Usability testing conducted as part of the United States Department of Education's National Institute on Disability and Rehabilitation Research (NIDRR) SBIR grant H133S080032 yielded important results about how people interact conversationally with computers.

A major objective of the grant is to understand the feasibility and usability of a natural-language product, particularly for persons with visual impairment. We created a prototype application: a conversational personal information manager named JotChat, using our patent-pending Tridbit natural language technology. JotChat enables users to enter and retrieve a wide variety of information about people, contact information, relationships, dates, shopping lists, etc. using everyday human communication. Usability testing with sighted and visually impaired users, which we conducted in two rounds, was critical to gaining an understanding of how people interact with conversational applications on computers.

1. Our users embraced English as a powerful way to interact with computers

We hypothesized that, because the majority of input was via keyboard and because JotChat is still a development prototype, users would be able to comply with the test, but not necessarily with great enthusiasm. We were pleased to discover that, even given these limitations, users responded enthusiastically to interacting with a computer using English. Their reactions indicated the experience was enjoyable and not stressful.

We expected visually impaired users to be most enthusiastic since they represent an under-served population and would find conversational interface uniquely suited to them. By contrast, we expected sighted users to be more attached to their conventional personal information managers. Instead we found that sighted and visually impaired users responded with similar enthusiasm. This, along with how readily our users imagined other uses of JotChat after experiencing this limited feature set, validates our long-term perspective that Tridbit technology will drive a wide range of applications for improving human interaction with computers and other devices.

2. Untrained users were able to quickly master JotChat's straightforward English interaction

The learning curve was rapid without requiring detailed instructions, coaching by test administrators, or references to a manual. Users learned how to best interact without being explicitly aware of their adjustment, much as we do when we meet new people. We are unaware of any other natural language application that has achieved this ease of interaction.

3. Speech input emerged as a viable option for JotChat interaction

We were pleased with the quality of the commercial speech recognition software engine we used. Its accuracy, while not perfect, did not overly interfere with the test flow or the user experience. We expected that user's preference for speech vs. keyboard would correlate to the speech recognition accuracy they experienced, but this was surprisingly not the case. Our speech users took delight in being able to truly converse with a computer, even if that conversation was occasionally limited by speech input accuracy.

We were also somewhat surprised to find that speech input did not significantly affect what users said to JotChat compared to that of our keyboard users for the same test scenarios. Taken together, these results justify research to develop a next generation conversational interface that integrates speech input with keyboard and other methods to provide the best interface for each user and each situation.

The testing also provided initial validation for our methodology of managing advances in applied natural language understanding through short iterative cycles of directed R&D followed by usability testing. Our first usability test provided a baseline as well as target sentences to guide the next round of R&D. The second usability test verified progress with improved performance on all measures and a new batch of target sentences.

Finally, we discovered that the usability testers can also serve as an informal, but experienced focus group to guide application development. A structured discussion following the formal test sequence revealed how users rated their likelihood of using JotChat for various functions and applications, including differences between what visually impaired vs. sighted users considered critical features.

The body of this report details the insights and results we obtained, which will guide our future work, developing JotChat into a product that could be revolutionary in what it can do for the visually impaired and guide the way toward a conversational interface paradigm for all computer users.

19 April 2009

Karen Blaedow, Principal Investigator
Custom Technology Ltd
karen@customtechnologyltd.com

and the JotChat team:

Neal Ewers
Kathy Ley
Matt Peterson

Section 1: Overview of Testing

Test Background

A natural language understanding technology that strives for practical application has to account for the wide variety of forms that people use to express the same concept. For example, in this test, 20 users responding to the same 26 scenarios came up with 419 unique inputs. Users regarded these as legitimate “natural” responses to the test scenarios. (It is worth reviewing Appendix B just to fully understand the scope of the problem and the significance of our results.) Thus, usability testing is critical to ensure that a natural language-based system understands a significant set of potential responses.

Usability testing for this SBIR project was executed in two rounds: a first test about midway through the project (Task A5.1) and a second toward the project’s end (Task A5.2). The first test allowed us to assess how users interacted with an earlier version of the JotChat prototype and to use those inputs not understood by JotChat to guide extensions to JotChat. The objective was to both improve its language understanding and its usability by the visually impaired for the second test. Both tests followed a similar design, although we extended the second test to include speech input in addition to keyboard input. When appropriate we will refer to the results of the first test, which can be viewed in detail at:

<http://www.tridbits.com/pubs/ConversInterfaceReport1.pdf>

The specifics of how the patent-pending Tridbit natural language technology works are not covered in this document. To learn more about the Tridbit technology that underlies JotChat see:

[Blaedow, K. 2007] *Babble: Simple Conversations With a Computer*. Proceedings of the 2007 Semantic Technology Conference, San Jose, CA. URL = <http://www.tridbits.com/pubs/simpleconvers.pdf>.

We anticipate a forthcoming paper to be published at the end of this project to describe further advances in Tridbit technology generated by this SBIR.

Description of the test procedure and subject demographics

Sighted and visually impaired users were selected from the local community to represent a mix of genders, ages, and computer abilities. In all there were 20 users: 10 sighted and 10 visually impaired. Eight of our visually impaired users were blind and 2 were low vision. The testing took place over a two-week period in March 2009.

After signing the consent form, users were read the introduction to the testing process and then presented with a practice question followed by the 26 scenarios. Appendix A contains the test script.

Most test sessions were conducted with one moderator who read the scenarios to the users, a note taker, and an observer. JotChat produced a log of each test session for later analysis.

An audio option allowed visually impaired users to hear their typed input echoed as well as hear synthesized audio of JotChat’s response. (Note that our visually impaired users were largely familiar with this form of computer interaction through the use of screen reader accessibility technology.)

Speech input was tested on a portion of the users for the final 8 scenarios of the test sequence. The remainder of the users served as a control group, continuing to use the keyboard.

Each scenario asked the user to either get information from JotChat or provide information to JotChat. Users were told to use simple English sentences to communicate with JotChat. For example scenario six asks the user:

What if you wanted to know all of Bob's phone numbers? How would you get this information?

The user responds by typing English sentences on the keyboard. Below is an example interaction:

Tester: Can I have all of Bob's phone numbers?
JotChat: I don't understand. Can you think of another way to say it?

Tester: What are all of Bob's phone numbers?
JotChat: 1. 234-5678 (work phone number);
2. 234-8765 (home phone number);

In this example, JotChat did not (at the time) understand the user's first response and asks the user to think of another way to say it. The user then types a question JotChat understands and JotChat provides the requested phone numbers.

Section 2: Overview of Test Results

The results and analysis of our usability testing are detailed in Sections 3-6. This section presents a quick overview of the key results and contents of each of the next four sections.

1. The major result, which was evident in the first usability test and confirmed here, is that our users found it easy and enjoyable to interact with JotChat. They continued to treat JotChat as a conversation partner, but possibly less human than the first round. These observations are categorized and discussed in Section 3, *Treating JotChat as a conversation partner*.
2. Performance, as defined by the ease with which users successfully completed each scenario, improved on all measures. We attribute these improvements mostly to enhancements to JotChat guided by the results of the first usability test. This provides initial validation our methodology of achieving advances in applied natural language understanding through short (~3-month) iterative cycles coupling directed R&D with usability testing. We plan to continue this approach in future work. Section 4, *Performance differences between tests and users* presents and discusses the test's performance results in full detail.
3. Speech input emerged as a viable option for JotChat interaction. We were pleased with the accuracy of the speech recognition engine we used, Dragon NaturallySpeaking® 10 ("Dragon"). We expected that user's preference for speech vs. keyboard would correlate to the speech recognition accuracy they experienced, but this was surprisingly not the case. We were also somewhat surprised to find that speech input did not significantly affect what users said to JotChat. The data from speech testing and a detailed discussion of these findings can be found in section 5, *Speech input findings*.
4. A structured discussion following the formal test sequence revealed how users rated their likelihood of using JotChat for various functions and applications, including differences between what visually impaired vs. sighted users considered critical features. Section 6, *Users reaction to JotChat*, organizes the content of these into data for analysis.

Overview of the data collected

Appendix B contains a table for each scenario listing all the unique responses for that scenario along with the number of times it was given and whether JotChat understood it. These tables record a total of 784 sentences or phrases that users input in response to the scenarios, of which JotChat understood 541, or a little over two thirds. However, two scenarios (16 and 17) ask users to do things with lists, a capability not yet developed within JotChat. (We included the list scenarios to find out how people would naturally ask about lists.) If those two scenarios are removed from the calculations, that leaves a total of 670 user inputs of which JotChat understood 503, or three quarters. The statistics in the remainder of the report will omit the list scenario questions, unless otherwise specified.

It is worth perusing these tables to see the many ways people come up with to communicate the same request. The average number of unique responses for a scenario is 17. Scenario 16 had the most variations listing 42 ways users tried to put milk on a list. Of the non-list scenarios, scenario 13 had the most variations listing 29 ways to ask who lives in Madison. Seeing all these language variations helps one appreciate what a complex problem it is to understand natural language.

It is also interesting to note that some scenarios seemed to naturally elicit more variation than others. For example there were many subtle variations and ellipsis for the topics dealing with phone numbers and relationships, but fewer when the topic of the scenario was how many children Kelly has (scenario 20) or Paul's nickname (scenario 9). Tridbit technology's model includes specialized language patterns within specific topic areas, allowing for this type of variation.

In addition to the raw responses and counts presented in Appendix B, Table A also provides a one line summary of the responses for each scenario.

Table A – Summary of Scenario Completion, Variation, and Understanding

The **Completed** column indicates the number of users who were able to come up with at least one sentence or phrase that fulfilled the scenario. The **1st try** column is the number who did it on their first try. The last two columns indicate how many unique inputs were entered and how many of those were understood.

Scenario	Completed	1 st try	Variations	Understood
1) You need to call your friend Paul but you don't know his phone number. What would you type to get this information from JotChat?	20	20	11	11
2) JotChat knows that Paul has a wife. How do you find who it is?	20	17	18	12
3) What if JotChat does not have the phone number of your friend, Alice. It is 221-4545. How would you enter this information?	20	15	14	5
4) Verify that JotChat now has Alice's phone number.	20	16	15	10
5) Bob works at a computer store. How would you ask JotChat for his number in order to contact him at work?	20	16	15	10
6) What if you wanted to know all of Bob's phone numbers? How would you get this information?	19	14	25	14
7) Your friend Paul's cell phone number is 222-3333. How would you give JotChat this information?	20	17	10	6
8) Bob's email address is bob@nomail.com. How would you enter this in JotChat?	20	17	13	8
9) You know Paul has a nickname but you can't remember it. Can you find this out from JotChat?	20	18	9	7
10a) A while back you told JotChat about Jim, but now you can't	20	15	23	8

remember who he is, how would you have JotChat jog your memory?				
10b) Which Jim? 1. Jim Eastman ; 2. Jim Rockford ;	20	15	14	6
11) How would you get Mary's address from JotChat?	20	18	13	9
12) You can also give JotChat addresses, but you need to put a quote at the beginning and end of the address. Also, JotChat will not yet recognize abbreviations, so completely spell out everything in the address. Given that, how would you enter Larry's address, which is: 111 Main Street, Madison, Wisconsin 53700	20	17	13	10
13) How would you ask JotChat to come up with names of people who live in Madison?	18	10	29	3
14) How would you ask JotChat for the company that Bob works for?	20	13	14	4
15) You'd like JotChat to give you a list of all the people you've entered who work at Cool Toys. What would you ask?	20	13	15	2
16) JotChat will be able to keep a list of things you need to do or get. If you wanted to have an item, say you are out of milk, appear on such a list, what would you tell JotChat?	17	6	42	11
17) How would you have JotChat display the list?	17	8	39	8
18) How would you ask JotChat for Larry's zip code?	20	18	9	7
19) What would you ask to get the address of Cool Toys?	20	16	13	7
20) How would you find out the number of children Kelly has?	20	19	6	4
21) How would you find out their names?	20	16	15	4
22) You have never met Bob's mother, but you need to call her. How would you get help from JotChat on this?	19	15	20	13
23) What would you ask to get Kelly's email address from JotChat?	20	16	9	6
24) If JotChat could place a phone call for you, how would you ask it to connect you with Bob?	20	20	7	7
25) Paul is having a birthday soon. Get the date from JotChat.	20	18	8	6

Section 3: Treating JotChat as a conversation partner

There is a continuum in the flexibility and sophistication of language-based computer human interaction. For example, "English-like" command languages include those offered by cable companies to control TVs, commands spoken to cell phones or even UNIX or SQL commands. One way to determine when the natural-language barrier has been crossed is to observe how users master and interact with a given language system. People have the ability and desire to communicate using natural language. Each individual finds it enjoyable rather than stressful when they can express a thought in his/her preferred way. On the other hand, people generally struggle to master "English-like" command languages, which lack the flexibility to allow people to express the same meaning in multiple ways.

Once again, our results show that people considered JotChat to be more than an "English-like" command language; however they treated it somewhat less human than in the first round. We cited 6 behaviors (listed below) as evidence of users treating JotChat as a conversation partner. All were still evident but a few were less strong. In addition to individual factors discussed below, one factor that likely contributed

to this was that the average user this round was younger, more tech-savvy and therefore less likely to humanize the computer.

In addition, we modified how we told users to think about JotChat. We continued to tell people to “use everyday language and not cryptic input” but were less emphatic about “talking to it just as you would another person.” After the first round we realized that natural language conversation with a computer is not exactly like another person, but more direct language without social undertones. We really didn’t need to explain this to people, after a few interactions they would adjust. In fact, it would be more confusing to explain than to just let them get the feel for it.

Finally the performance of JotChat in this second round improved so that users had to make fewer adjustments. JotChat understood three quarters of the users’ inputs as opposed to two thirds in round 1. 81% of the users’ first tries were understood by JotChat.

The nature of the sentences JotChat did and did not understand comes into play as a factor in allowing users to accept JotChat as a conversation partner. In general, JotChat could understand variations of an input that are hard for humans to discriminate. In other words it understands “The phone number of Paul is 123-4567” and “The phone number for Paul is 123-4567.” It would be very difficult for people to remember that only one or the other was valid.

Other variations for expressing the relationship between Paul and his phone number without changing the underlying meaning are also hard to discriminate, for example, “Paul’s phone number is 123-4567” or “123-4567 is the phone number for Paul.” JotChat recognizes many of the forms used to express these basic relationships so the user is not limited to using one specific way. It may even be the case that the underlying model is human-like in the way it accomplishes this, but the important thing for this discussion is that users sense that JotChat has human-like flexibility in what it understands and does not require cognitively difficult discriminations of them.

Users seemed to pick up on this consistent but flexible language style and treat JotChat as a conversation partner. Evidence that users considered JotChat a conversation partner includes:

- paying attention and adjusting to the capabilities of their conversation partner
- not wanting to go on until their conversation partner understood them
- being polite with their conversation partner
- not taking advantage of computer-style input options
- not taking advantage of the transcript box to look back at what had been said previously
- having a positive, non-stressed attitude toward the interaction

Paying attention and adjusting for their conversation partner

Humans naturally pay attention and adjust to the capabilities of their conversation partner. If their conversation partner has limits, such as an adult talking to small children or a native speaker talking to someone learning the language, the more capable speaker will adjust their language to what their partner can understand. We do this instinctively.

Once again people quickly trained themselves to talk to JotChat. It was a little harder to observe this round mainly because users did well right off the bat. For example, all 20 users came up with an understood response to the first scenario on their first try! With such a strong start, it was harder to detect improvement, but it was there.

Users in the first round started with more language patterns that JotChat did not understand, especially indirect language such as “Could you give me . . .”, “Do you have. . .” or “How would I find. . .” These were quickly extinguished and hardly used by the end of the test.

Second round users presented less indirect language, but more abbreviated language, especially when users were asked to input data such as phone numbers or addresses. The first time users are asked to input a phone number in scenario 3, there were 3 non-understood responses that were more computer commands than sentences as in “Alice number: 2545643”. In addition, 4 inputs failed to provide the phone number. Both these styles of inputs were quickly adjusted to something JotChat would understand.

Wanting their conversation partner to understand them

If the user’s first couple sentences were not understood by JotChat, we let the user dictate whether they wanted to continue to try other ways of phrasing their request. Again we saw reluctance on the user’s part to go on without figuring out a sentence JotChat would understand. Of the 480 non-list scenarios the users attempted, they abandoned only 4 of them without figuring out a sentence JotChat would understand. This is down from 19 abandoned scenarios in round 1.

One of the characteristics about the system that encouraged this behavior was the user’s belief that they could find a way to communicate. Not just any way, which is also true of English-like command languages, but a way that “made sense” to them and would help them in subsequent scenarios. Thus the discovery that direct questions like “What is Paul’s number?” work better than indirect questions such as “Can you give me Paul’s number?” can be processed easily by the user to adjust their interaction with JotChat. If it were the case that JotChat understood “the number of Paul” but not “the number for Paul” (it understands both) this would be more difficult for humans to keep straight since the two expressions seem the same.

Being polite with their conversation partner

“Please” was used 21 times, similar to the 22 times last round. That certainly indicates the users were not treating JotChat like an ordinary computer program. It was common for people in round 1 to use the more respectful and polite ways of indirectly asking for information such as “Could you tell me . . .”, “Do you have. . .”, “Let me give you...” We added the ability for JotChat to understand some indirect language, but users this round used only a handful of indirect requests, for example “May I have Mary’s address?” Given the small sample of users, it is not surprising to see this kind of variation in speaking styles from round 1 to round 2, especially given the younger more tech-savvy character of the second round users.

Not taking advantage of computer-style input options

Scenario 10 was constructed to create a situation where the users would have to resolve an ambiguous reference. The scenario asks them to “have JotChat jog your memory” about a person named Jim. It turns out JotChat is aware of two people named Jim and asks the user to resolve this ambiguity by giving them the following choice:

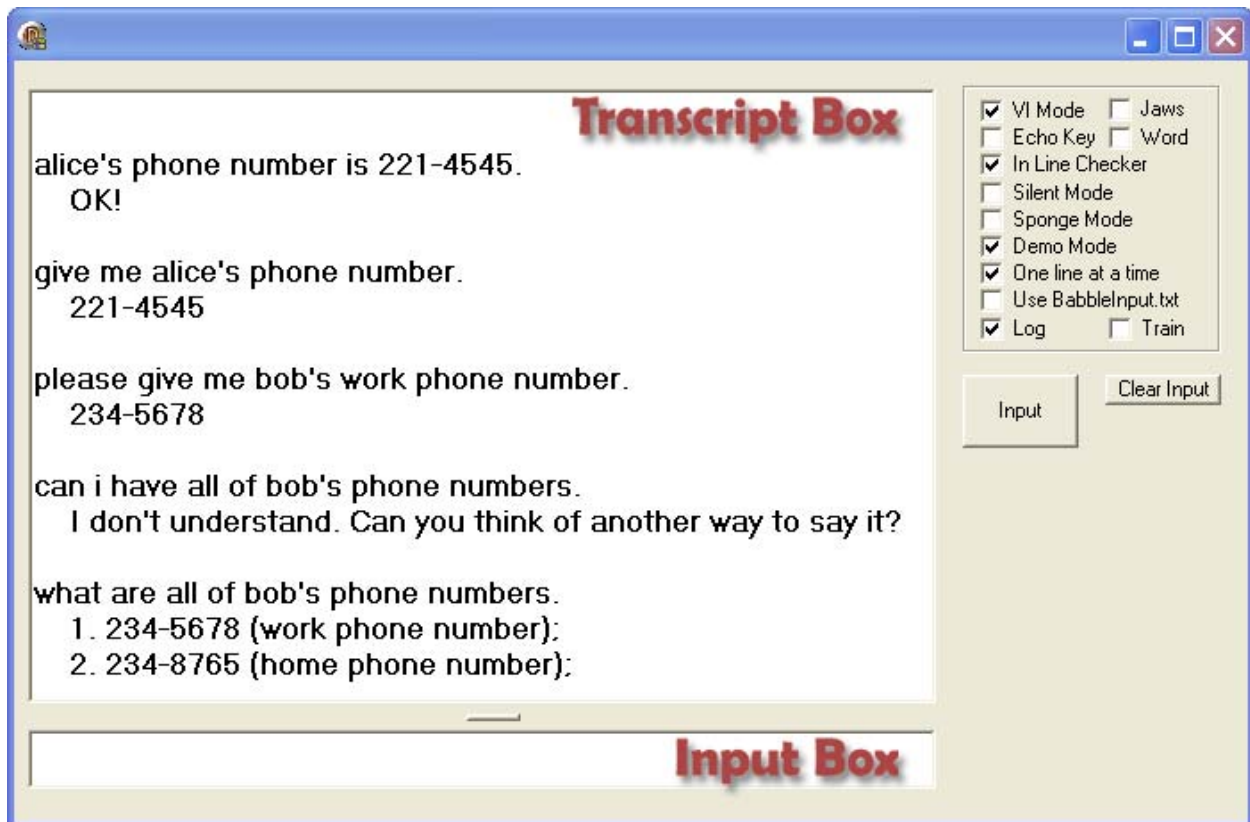
Which Jim?

1. Jim Eastman ;
2. Jim Rockford ;

In round one, only 4 of the 19 users thought to respond by typing a 1 or 2. When users were asked why they didn't type a number, they generally said that it didn't occur to them because they were in conversation mode.

In round two, 13 of 20 users responded by typing a 1 or 2. Once again this round of users seems more inclined to treat JotChat as a computer, with the added capability of understanding English. This is a good thing as the intent of providing a numbered list was to give users a shortcut for making their choice. But it also indicates that more work needs to be done to ensure that JotChat initiates requests of the users so that the user is not confused as to how to respond.

Not looking back at the transcript box



The JotChat interface, shown above, has an input box where the user types and a transcript box that has a record of the ongoing conversation. For visually impaired persons there is an accessibility interface that provides a great deal of control in navigating and reading this transcript box. Given the short time frame for the tests one would not expect the visually impaired users to take advantage of this information, although we did explain the interface to several of the visually impaired users who asked. The interesting thing to note, however, was that the sighted users did not appear to make much use of this information either.

One would expect that if the language was difficult to come up with, as might be the case with a command language like UNIX, one would look back in the transcript box to see what worked previously and type it again. Instead all users seemed to operate conversationally, where they simply used their natural ability to remember what was said previously, without needing to refer back to a transcript. When given the choice to edit what they had said or to say it again, most users chose the conversational approach and asked the question again in a different way.

None of the scenarios required specific information from previous scenarios, but there was sufficient context from one scenario to the next to create opportunities to use pronouns. While JotChat is capable of handling singular pronouns, few users made use of this. Those that figured it out (we didn't tell users) generally liked the idea and continued to use pronouns. Unfortunately, the pronoun most likely to be used was "their", asking "What are their names" to get Kelly's children's names in scenario 21. "Their" is the only pronoun not implemented at the time of these tests.

Having a positive, non-stressed attitude toward the interaction

In the follow up questions asking users what they liked and disliked, not a single tester in either round labeled the exercise as tedious. Given that most responses were entered via keyboard and the speech recognition software required at least 20 minutes of setup and training, it would not have surprised us to have a few testers find that tedious. Instead every single tester responded positively in terms of enjoying the interaction. They said it was fun, that they would like to interact with computers in this way and many wanted to be repeat testers.

In addition to these documented comments is the subjective observation that many of our users were somewhat apprehensive at the start of the test but after a few scenarios relaxed and became very comfortable interacting with JotChat. This seemed especially true of those who were computer novices as well as the visually impaired users who understandably assume that learning any new computer program would be a major undertaking.

Should we really be surprised?

We saw a definite excitement about using JotChat technology. Perhaps we shouldn't be surprised. For 30 years people have been adapting their brains to how a personal computer wants information. With JotChat, our users could see that finally, even in its primitive form, they could give information to a computer in a form that was comfortable to them and get it back in a very natural way. Our users picked up that this was revolutionary, and they liked it. In fact, the only dislike consistently mentioned by users was the desire to have JotChat understand more ways of saying things. We are working on that.

Section 4: Performance differences between tests and users

Users' performance improved on all measures between round 1 and round 2. Table B on the following pages shows the completion, variation, and understanding measures for all the scenarios in the second round compared to matching scenarios in the first round. In addition, the table on the far right shows the relative difference in these measurements between the 6 users that participated in both rounds, i.e. the repeat testers, and round 2 averages.

The main thing to take away from this analysis is the improvement in JotChat's ability to understand user responses. This was most pronounced in the 18 scenarios that were shared between the two rounds of testing. In round 2, users were able to come up with a request to satisfy the scenario 99.4% of the time.

Only twice did a user go on without completing the scenario. In round 1 this happened 5 times more often for a total of 10 incomplete scenarios. That is still a very impressive round 1 score of 97% of the time that a completely untrained user came up with a natural language request that JotChat understood and was able to process in order to satisfy the scenario. Further, in round 2, 84% of the users' first attempts were understood as opposed to 71% in round 1. Because users had more of their first tries understood they produced fewer variations, of which a higher percentage were understood.

Successful method for improving JotChat comprehension

These results show that our method of using non-understood results from previous usability tests to feed into extension sets worked well to expand JotChat's comprehension in our targeted area.

These enhancements are generally not limited to the specific examples being worked on. Improvements tend to spill-over to other language constructs, especially when filling in or enhancing the underlying tridbit technology model that describes meaning structures and how meaning is produced from surface structure of natural language.

This spill-over effect is supported by the fact that the average performance scores across all 26 scenarios in round 2, including new scenarios, were better than the averages for all scenarios in round 1. The scores for completed scenarios and completed on 1st try for round 2 overall were 98.1% and 77.5% including list questions and 99.2% and 81.0% not including list questions. The overall completed scenarios and completed on 1st try for round 1 was 95.4% and 70.4%

Repeat testers perform better

Some of the improvement could be attributed to the 6 repeat testers from round 1 whose performance exceeded the group average. However the difference between the repeat testers and the round 2 group averages was far less than the difference between round 1 and round 2 group averages as evidenced by comparing the 2 rightmost tables on the following page.

It is worth noting that the performance boost in the repeat testers was equal for both the shared scenarios and the new ones. This suggests that these users retained a generalized notion of a conversational style to use with JotChat rather than any memory of the specific scenarios. This is corroborated by one of the repeat testers who had a tendency to use indirect language in the first round. The tester adjusted their interaction such that by the end of round 1 they were using the more direct language that works best with JotChat. In the second usability test, over three months later, this tester did not use any indirect language with JotChat.

Obviously one must use caution drawing conclusions from such small samples. Nonetheless, every measurement, not to mention common sense, points to humans having facilities with natural language even beyond understanding it.

Little difference between sighted and visually impaired users

Overall, there really was very little difference in how sighted and visually impaired users performed. One visually impaired user warned us in round 1 that if the test took sighted people 45 minutes, it would probably take at least double that for a visually impaired tester. In fact that tester finished faster than the previous sighted tester. Conversation is a mode of interaction that does not significantly favor sighted users.

Sighted users do have the advantage of seeing the screen, but as stated in the previous section, they did not appear to make use of it. Visually impaired users must rely on the auditory output from JotChat. The accessibility interface allows some choice in the amount of auditory feedback ranging from echoing each keystroke to just echoing the sentence the user types followed by JotChat's response. We upgraded the default voice we used in round 2 from Microsoft Mike to Cepstral David. A tradeoff of the better quality sampled voices is that they have more lag so do not work as well for key echo. When users wanted this feature we switched back to Microsoft Mike.

Measurement issues

It is also worth noting some of the subtleties and imprecision in the measurements we are using. It is problematic to measure performance on a task where we have told users there is no right and wrong, that their non-understood responses are at least as valuable as their understood responses in that they help us expand JotChat's knowledge. The users' ability to come up with at least one response that satisfies a scenario provides a very basic measurement of JotChat's performance.

We also used how often the users' first response is understood as a more sensitive measure of JotChat's performance. After the first response the tester may become influenced by JotChat's response or by discussion with the test administrators, so the first response is also the least biased. (We encouraged multiple responses where the user was so inclined in order to gain more information for future JotChat development.)

There were sequences of responses where it was unlikely the user's first try was their "best attempt." Users in round 2 sometimes experimented with JotChat's understanding by inputting a sequence of responses such as the following sequence for scenario 3 which asks the user to enter the phone number for Alice:

Tester: Alice number:2545643.
JotChat: I don't understand. Can you think of another way to say it?

Tester: Alice number is 2214545.
JotChat: I don't understand. Can you think of another way to say it?

Tester: Alice's phone number is 2214545.
JotChat: OK!

Table B – Performance comparison between rounds and testers

The leftmost columns list completion, variation, and understanding scores for each scenario in round 2 as percentages. Scores for corresponding scenarios in round 1 are lined up in the following columns for comparison. The next series of columns calculate the differences between round 2 vs. round 1 scores. The last series of columns calculates the differences between repeat testers vs. average round 2 scores. The three lines at the bottom of the table list averages across various sets of scenarios.

Scenario	Completed	1 st try	Variations	Understood	Scenario	Completed	1 st try	Variations	Understood	Completed	1 st try	Variations	Understood	Completed	1 st try	Variations	Understood
	Round 2	Round 2	Round 2	Round 2		Round 1	Round 1	Round 1	Round 1								
Round 2					Round 1												
1) You need to call your friend Paul but you don't know his phone number. What would you type to get this information from JotChat?	100%	100%	0.55	100%	1) Same	100%	74%	1.00	63%	Max	26%	-0.45	37%	0%	0%	0.12	0%
2) JotChat knows that Paul has a wife. How do you find who it is?	100%	85%	0.90	67%	2) Same	100%	84%	0.84	63%	Max	1%	0.06	4%	0%	15%	-0.23	33%
3) What if JotChat does not have the phone number of your friend, Alice. It is 221-4545. How would you enter this information?	100%	75%	0.70	36%	3a) Same	100%	63%	1.00	37%	Max	12%	-0.30	-1%	0%	8%	-0.03	39%
4) Verify that JotChat now has Alice's phone number.	100%	80%	0.75	67%	3b) Same	100%	89%	0.58	73%	Max	-9%	0.17	-6%	0%	-13%	0.42	-10%
5) Bob works at a computer store. How would you ask JotChat for his number in order to contact him at work?	100%	80%	0.75	67%	5) Substituted Bob for Joe	89%	63%	1.11	38%	11%	17%	-0.36	29%	0%	3%	0.08	13%
6) What if you wanted to know all of Bob's phone numbers? How would you get this information?	95%	70%	1.25	56%	6) Substituted Bob for Joe	84%	37%	1.63	23%	11%	33%	-0.38	33%	5%	13%	-0.25	27%
7) Your friend Paul's cell phone number is 222-3333. How would you give JotChat this information?	100%	85%	0.50	60%	7) Same	95%	84%	0.79	40%	5%	1%	-0.29	20%	0%	15%	0.00	40%
8) Bob's email address is bob@nomail.com. How would you enter this in JotChat?	100%	85%	0.65	62%	14) Substituted Bob for Joe	95%	74%	0.47	56%	5%	11%	0.18	6%	0%	-18%	0.85	-28%
9) You know Paul has a nickname but you can't remember it. Can you find this out from JotChat?	100%	90%	0.45	78%	16) Same	100%	84%	0.47	67%	Max	6%	-0.02	11%	0%	10%	0.22	22%
10a) A while back you told JotChat about Jim, but now you can't remember who he is, how would you have JotChat jog your memory?	100%	75%	1.15	35%	22a) There is someone named Jim, you want to find some information about him. How would you begin?	100%	47%	0.79	20%	Max	28%	0.36	15%	0%	-8%	0.18	3%
10b) Which Jim? 1. Jim Eastman ; 2. Jim Rockford ;	100%	75%	0.70	43%	22b) Same	95%	95%	0.37	71%	5%	-20%	0.33	-29%	0%	8%	0.30	40%
11) How would you get Mary's address from JotChat?	100%	90%	0.65	69%	19) Substituted Mary for Linda	100%	89%	0.63	75%	Max	1%	0.02	-6%	0%	-7%	1.35	6%

12) You can also give JotChat addresses, but you need to put a quote at the beginning and end of the address. Also, JotChat will not yet recognize abbreviations, so completely spell out everything in the address. Given that, how would you enter Larry's address, which is: 111 Main Street, Madison, Wisconsin 53700	100%	85%	0.65	77%	New									0%	-2%	0.35	6%
13) How would you ask JotChat to come up with names of people who live in Madison?	90%	50%	1.45	10%	New									10%	17%	-0.62	10%
14) How would you ask JotChat for the company that Bob works for?	100%	65%	0.70	29%	New									0%	2%	-0.03	21%
15) You'd like JotChat to give you a list of all the people you've entered who work at Cool Toys. What would you ask?	100%	65%	0.75	13%	New									0%	35%	-0.58	87%
16) JotChat will be able to keep a list of things you need to do or get. If you wanted to have an item, say you are out of milk, appear on such a list, what would you tell JotChat?	85%	30%	2.10	26%	New									15%	-30%	0.07	12%
17) How would you have JotChat display the list?	85%	40%	1.95	21%	New									-2%	-40%	1.05	2%
18) How would you ask JotChat for Larry's zip code?	100%	90%	0.45	78%	New									0%	-7%	-0.12	-28%
19) What would you ask to get the address of Cool Toys?	100%	80%	0.65	54%	New									0%	-13%	0.68	21%
20) How would you find out the number of children Kelly has?	100%	95%	0.30	67%	9) Substituted Kelly for Jane	100%	95%	0.37	86%	Max	0%	-0.07	-19%	0%	5%	0.20	33%
21) How would you find out their names?	100%	80%	0.75	27%	10) Same	100%	53%	1.05	45%	Max	27%	-0.30	-18%	0%	20%	-0.08	73%
22) You have never met Bob's mother, but you need to call her. How would you get help from JotChat on this?	95%	75%	1.00	65%	12) Substituted Bob for Joe	100%	79%	0.89	71%	-5%	-4%	0.11	-6%	5%	8%	0.00	18%
23) What would you ask to get Kelly's email address from JotChat?	100%	80%	0.45	67%	13) What would you ask to get Paul's sister Jane's email address from JotChat?	95%	53%	1.37	42%	5%	27%	-0.92	24%	0%	3%	0.22	8%
24) If JotChat could place a phone call for you, how would you ask it to connect you with Bob?	100%	100%	0.35	100%	18) Substituted Bob for Joe	100%	63%	1.26	50%	Max	37%	-0.91	50%	0%	0%	0.15	0%
25) Paul is having a birthday soon. Get the date from JotChat.	100%	90%	0.40	75%	21) Same	95%	47%	1.00	21%	5%	43%	-0.60	54%	0%	-7%	0.10	-8%
Averages for 26 round 2 scenarios	98%	78%	0.81	56%										1%	1%	0.17	17%
Averages for 24 non-list scenarios	99%	81%	0.70	58%										1%	4%	0.14	18%
Averages for 18 shared scenarios	99%	84%	0.68	63%		97%	71%	0.87	52%	2%	13%	-0.19	11%	1%	3%	0.20	17%

Section 5: Speech input findings

14 of the round 2 users provided responses to scenarios 18 through 25 using speech input. The other 6 users served as a control group, continuing to use the keyboard. Dragon NaturallySpeaking 10 was used as the speech recognition engine.

Speech recognition accuracy

Dragon's accuracy was better than expected, especially when we had the users go through Dragon's general training, which is described below. We tried skipping this training with our first user, but it had a more negative effect on performance than prior experimentation suggested. Thus we did the 20-30 minute training for the remainder of the speech users.

Dragon accurately transcribed 117 of the 149 responses spoken by users for an accuracy rate of 79%. If we only count the trained users, the accuracy jumps to 110 of 128 responses for an accuracy rate of 86%. Note that this accuracy rate reflects getting an entire response correct rather than the word accuracy rates often cited. Since responses were between 2 and 12 words long, the word accuracy rate would be much higher. But to be useful for natural language understanding at this point in time, the entire sentence must be accurately transcribed.

Selecting the appropriate microphone

Our research and our own particular ease of use needs suggested that certain microphones might yield better results than others. First, the microphone needed to be designed for speech with appropriate noise canceling circuitry. Secondly, it had to be one that visually impaired individuals could easily place on their heads. Thirdly, the microphone boom needed to be one that could be positioned well below the mouth to keep down popping and blowing noises. Fourthly, it had to have a mute button that could be easily used by the subject. Our final choice was the VXI TalkPro USB 100 7.02.

Training users for Dragon speech input

We used the Dragon suggested method of training which included having people first read two short text excerpts while Dragon checked voice quality and volume. Subjects then read one of the Dragon preferred training passages in order for Dragon to better understand their voice. Subjects with vision read the information as it was presented on the screen. Visually impaired subjects simply counted during the short quality and volume tests. These subjects completed the longer voice test by reading a Braille transcription of the screen text. In one case where the subject did not know Braille, information was spoken to the subject, phrase by phrase, and the subject repeated each phrase into the microphone.

Subjects were then asked to turn on the microphone and speak their response to the given scenario instead of typing it. They then turned off the microphone so that the input from the speaker would not be heard by Dragon. Before the subject's response was sent to JotChat we checked the information in the input field to see if Dragon had accurately recognized what was said.

We noticed that some people seemed to speak more clearly than others: Their words were better formed with fewer slurred syllables. However, the speech input test showed that Dragon recognized their speech patterns just as accurately as those whose speech seemed to us to be clearer. This is obviously quite subjective, but it would suggest that speech recognition may be more accurate for a wider range of speakers than we had at first thought.

Preference for speech vs. keyboard

We asked the subjects who used speech input four additional questions about their experience with speech input. The third question asked them which they preferred for input, speech or keyboard. The question was carefully worded to instruct the user to base their choice on the performance of the speech recognition they just experienced, not a conjecture of how it might improve or be different in the future. The expectation was that the user’s preference would be heavily influenced by the how well their speech was recognized. Surprisingly this was not the case. Table C below has a line for each user that did speech input that shows their stated preference next to measures of how well Dragon recognized their speech and their demographic data. While there was a weak correlation between speech recognition performance and the user’s preference, the user’s demographics was a better predictor of the user’s preference for speech. Specifically, visually impaired users were reluctant to give up their keyboards.

Table C – Speech preference compared to accuracy and demographics

The column 1 indicates the user’s preference for speech vs. keyboard. Column 2 indicates whether users thought speech input would increase their use of JotChat. Columns 3-12 are various measures of how well Dragon recognized their speech and demographic data which is described in the Key column of Table D below.

Preference	Affect Usage	# of Inputs	Derrs	%Derrs	Had to Key	VI	Age	1st Try Score	Sex	Skill	Repeat
Speech	+	13	0	0%	0	N	61	18	M	2	N
Speech	+	15	1	6%	0	N	41	15	M	3	N
Speech	+	9	1	10%	0	N	61	21	F	2	Y
Speech	+	10	2	17%	1	N	21	18	M	3	N
Speech	+	9	2	18%	1	N	61	18	M	2	Y
Speech	+	9	2	18%	1	N	41	24	F	3	N
Speech	+	8	0	0%	0	Y	41	20	F	2	Y
Keyboard	0	8	0	0%	0	Y	41	23	F	3	N
Keyboard	0	8	1	11%	0	Y	41	23	M	3	N
Keyboard	0	8	2	20%	0	Y	21	22	F	3	N
Keyboard	+	9	3	25%	0	Y	41	20	F	2	Y
Keyboard	0	8	4	33%	1	Y	21	23	F	3	N
Keyboard	0	11	14	56%	4	Y	21	19	F	3	N

Table D – Correlations between speech preference, accuracy and demographics

Variable	Correlation	Key
VI	-0.86	Is the user visually impaired?
Age	0.54	User’s age: 21-40, 41-60, 61 and better
1st Try Score	-0.49	Number of scenarios where the user’s first response was understood by JotChat
%Derrs	-0.48	Percentage of the user’s speech inputs that were not accurately transcribed by Dragon
# of Inputs	0.42	Total number of responses to the 8 speech scenarios given by this user
Sex	-0.41	User’s sex
Skill	-0.41	User’s computer skill level: 1=low, 2=medium, 3=high
Derrs	-0.40	Number of inputs the user spoke that were not accurately transcribed by Dragon
Repeat	0.28	Did the user participate in the first round of usability tests?
Had to Key	-0.19	Number of times the user had to key in their response because Dragon did not transcribe

Multiple things are likely going on here. There certainly is a minimum accuracy level below which users would not prefer speech. Dragon's performance for most users was above that level. Dragon's worst performance where users still said they'd prefer speech was 18%. Once performance sinks to missing 1 or more sentences out of 5, the users prefer keyboard. Only 3 users had an error rate above 20%. The lowest error rate was for our first tester where we tried going without training. All other users did Dragon's accuracy training, which significantly improved its performance.

Just meeting a minimum accuracy level did not guarantee a user would prefer speech. One tester who had all their speech transcribed perfectly, still preferred keyboard. All but one of the visually impaired users did not want to switch to speech input if that meant giving up their keyboards. While we didn't query the visually impaired users about this, there may be a couple reasons why many would find it hard to completely give up their keyboard for even a very good speech interface.

The visually impaired users may have been concerned with unanswered questions about how they would perform basic functions if the interface were all-speech. For example how could they verify what Dragon transcribed, edit a response or examine the information returned to them by JotChat? Sighted users may not have been as concerned because seeing the screen could help and even eliminate the need for some of these tasks such as seeing what Dragon transcribed and seeing what is in the transcript box.

But even sighted users might need to scroll a window or not even have windows in an all-speech deployment. Sighted users can assume these interface details will be worked out in some manner they will be able to use while visually impaired users have learned not to make that assumption. A visually impaired user may be wisely cautious in not giving up the familiar and proven interface of a keyboard for even the exciting possibility of speech input until they are certain it will work for them.

In a similar vein, our visually impaired users were well aware of the unreliability of speech recognition, some having tried it to control devices in the past. While Dragon's performance was remarkably good, something which nearly all users remarked on and were delighted to discover, users knew there would still be times when it would be difficult or impossible to have their speech accurately transcribed. For a visually impaired user who is thinking that JotChat could become an indispensable tool in helping to organize their life, it is more important that the interface be reliable and predictable than convenient but occasionally unusable.

By contrast, the sighted users are probably not seeing JotChat as an indispensable tool. While they found the experience of interacting with the computer in English very positive, alternatives for sighted users to keep track of this type of information are many. So the convenience factor for sighted users is actually tipping the evaluation of whether they would use JotChat as opposed to another method.

This finding is further supported by the answer to evaluation question 8, "If JotChat were able to use speech input, how would that affect the tasks you might use it for?" The column in Table C above labeled "Affect usage" distills the users response down to either a + if it would increase their usage or 0 if they'd use it about the same. Every sighted user answered that it would increase their usage, where as only 2 of the 7 visually impaired users said it would increase their usage. The basic message from our visually impaired users was that if JotChat were available (many asked when it would be), they'd use either keyboard or speech to have this kind of accessible functionality.

The message for future development is to continue to integrate speech, with dual motivations for the two groups. Speech is needed to entice sighted users into using JotChat to replace some of their current

methods for keeping track of personal information. This expands our potential market and gives us the opportunity for truly universal design. While our visually impaired users were satisfied with keyboard input, the idea that speech input might actually be viable was probably more exciting to them than the sighted users, although everyone was wowed by the experience. One user told us, “It made me feel like Captain Kirk.” Our goal should be to develop a next generation conversational interface that integrates speech input with keyboard and other methods that would provide the best interface for each user and each situation.

Speech does not affect users responses

Another surprising result was how little effect speech input had on what the users said. The tables in Appendix B have a keyboard symbol indicating which inputs were entered via keyboard for scenarios 18-25. Browsing through these tables shows little difference between the keyboard and speech responses. This contradicts what we saw last round with the phone testers who had a tendency to use more complex speech patterns. It also contradicts the common belief that when people use spoken language their speech will always be sloppier. The most likely explanation for this is that testers switched to speech after doing 17 scenarios on the keyboard, so were already set into a particular manner of speaking. It would be interesting to see what happens if people use speech input from the start.

On the other hand we were presented with an ideal result. Despite some claims to the contrary, we’ve shown that people are absolutely capable of inputting sentences that are as clean and direct as what they would type on a keyboard. In many cases the speech input was cleaner because Dragon was a better speller and better at putting apostrophes in the right place than most users. The exception was when users responded with “What is Cool Toys’ address?” to scenario 19, which asks users to get the address of Cool Toys. In this case Dragon left out the apostrophe. Since this is one of cases where JotChat needs the apostrophe, we inserted the apostrophe for Dragon. Eventually this problem will be solved by JotChat not requiring apostrophes as it currently does not require capitalization or end of sentence punctuation.

What users liked and disliked about speech

Not having to deal with spelling and punctuation was one of the things users mentioned when asked what they liked about using speech with JotChat in evaluation question 5. They also said they were “impressed with how easy it is” and “it was very user friendly. I was impressed with its precision, its accuracy. I liked its quick feedback.” They liked that it was a fast, direct form of communication that freed their hands to do other things. In short, they thought it was easy, fun, quick, and were pleased that it understood them and their questions.

When asked what they did not like about using speech with JotChat: Some testers experienced frustration when Dragon didn’t understand what they said. Testers wanted the ability to verify their speech was accurately transcribed. They questioned how it would work in noisy environment and the inconvenience of using a microphone.

It is hard to downplay the significance that having reliable speech recognition would make to the application of natural language technology such as Tridbits. We were pleased to discover that with the latest commercial software, the current state of speech recognition is finally approaching the point where it could satisfy a significant set of users, at least under good acoustical conditions. But our most exciting discovery is how much better speech recognition becomes when JotChat natural language capability is used to improve the accuracy of the speech recognition front end. In Task B3, we achieved a significant

accuracy improvement when we integrated JotChat with speech recognition software's assessment of candidate utterance interpretations and anticipate additional improvements when we integrate with the software's training capabilities. Task B3 research and future potential is detailed in a report available at: <http://www.tridbits.com/pubs/SpeechRecEnhance.pdf>

Section 6: Users' reaction to JotChat

In round 1 we asked users what other kinds of information or tasks they would like JotChat to handle. Users came up with an extensive list of ideas which we boiled down to 18 in order to have users rank which tasks they would be most likely to use. The table below shows the average score for each task in order of popularity across all users. Average scores for the visually impaired users are given in the rightmost column and is an average of 0.14 points higher.

Table E – User rankings of Tasks JotChat might perform in the future

Scoring: 2=Definitely WOULD use; 1=MIGHT use; 0=Would definitely NOT use

Tasks JotChat might perform in the future	All	VI
1. Enter & retrieve names, addresses, phone #s, and personal info	1.80	1.90
2. Get address and other info for stores, restaurants, theatres, via MapQuest, on-line phone books or similar service	1.80	1.90
3. Enter and retrieve date & times of appointments, birthdays, due bills, etc.	1.75	1.80
4. Conversational interface to general web information sources such as dictionaries, encyclopedias, or Wikipedia	1.70	2.00
5. Secret keep (credit card numbers, social security numbers, passwords)	1.65	1.70
6. Bar code reader for grocery stores	1.65	2.00
7. Conversational weather queries and reports	1.60	1.70
8. Appliance front end	1.55	1.90
9. Maintain lists of books and other media including interface to library catalogs, iTunes, etc.	1.50	1.70
10. Print address labels	1.45	1.30
11. Set up reminders about appointments, birthdays, due bills, etc.	1.45	1.80
12. Create and maintain shopping and to do lists	1.45	1.60
13. TV and radio listings	1.30	1.50
14. Conversational interface to bring up new or previously identified web pages (Go to Trace's website)	1.30	1.30
15. Bus schedules	1.15	1.70
16. Synch with Microsoft Outlook	1.05	1.00
17. Recipes	1.05	1.10
18. Family tree keeper	0.90	0.80

The top three ranked tasks are the basic functions we expect to incorporate into JotChat, namely entering and retrieving information for personal contacts, retrieving publically available business and contact information and entering and retrieving dates and appointments. While these were highly rated among visually impaired users, there were four other functions that rated just as high or higher.

Conversational interface to general web information

The fourth ranked task for the group, but tied for number one for visually impaired users was a conversational interface to general web information sources such as dictionaries, encyclopedias, or

Wikipedia. Who wouldn't want a conversational way to access any information on the web and get back just the desired information rather than pages of links?

For visually impaired users, navigating a set of links to find a specific piece of information is challenging. As long as many web sites remain either completely or partially inaccessible to screen readers, searching for information in an encyclopedia or Wikipedia, for example, is out of the question. And many simply lack either the web savvy or the actual computer to search for online information.

Unfortunately, making JotChat capable enough to understand and retrieve arbitrary web content is futuristic, but a long-range possibility. Existing efforts to access online information in conversationally friendly ways use clever search strategies that work for limited types of questions or even behind the scenes human agents. While this is too ambitious for a near-term goal, we may be able to offer specific functions like dictionary definitions, etc. (An interface to the soon on-line version of the Dictionary of American Regional English, Medical, or of Oxford's Dictionary would pull the massive power of the English language to any user.)

Bar code reader

The other number one ranked task for visually impaired was a bar code reader for grocery stores. Several of our users pointed out they would also use such a feature to identify the contents of their cupboards by simply reading cans and boxes with the bar code reader. "No more mystery dinner" commented one user.

Besides identifying an object, JotChat could pull up product details such as cooking instructions, ingredients, nutritional information, package size, storage suggestions, allergens, recalls, and food safety guidelines which are normally obtained only with the help of a sighted person. And if the user enters their acquisitions into JotChat's database, they could simply query JotChat about what they had rather than manually looking for it each time. Users could also add their own information about a product, including telling JotChat about allergies so it would warn them anytime they scan an object containing that allergen.

Implementation would require specific hardware and access to a database of bar code information. It may be possible to use a built in camera to do bar code scanning. Portable readers can currently be attached to some devices, like Palm Pilots. Several users were familiar with these units but commented on their cost.

While careful research needs to be done, bar code scanning capability should be an important consideration when evaluating handheld devices for future JotChat deployment.

Appliance front end

Another highly ranked task by our visually impaired users was using JotChat as an appliance front end, in other words to be able to talk to their appliances. People with various disabilities, including limited mobility and inability to see visual displays would suddenly find it possible to talk to their thermostats, stoves, and other appliances to do what most people take for granted, setting the temperature in their homes, for cooking or setting a wash cycle for their clothes.

Unfortunately this is not a task that we can accomplish on our own. It requires appliance manufacturers to build in standards that would make this type of communication possible. Some Universal Remote Control (URC) standardization has already been achieved through the work of the University of Wisconsin, Madison Trace R & D Center. We have partnered with Trace in the past to experiment with using JotChat as a front end to the standards they developed. We would welcome the opportunity to continue this work

when such standards become adapted by manufacturers. Even if one or two manufacturers or one or two types of appliances could be made to operate based on speech or keyboard input, the marketability of JotChat would be greatly enhanced along with the freedom of people who cannot easily accomplish these tasks on their own.

Set up reminders

The ability to set up reminders about appointments, birthdays, due bills, etc was asked as a separate task from entering and retrieving appointments. Our visually impaired users ranked setting reminders equal to entering the information while the sighted users ranked it significantly lower.

One explanation may be that people who have vision have enumerable ways of keeping this kind of information. However, putting a note on your refrigerator or placing an item at the door to remind you to take it, are not strategies that work for people who are visually impaired. A simple, conversational way to set and retrieve appointment information and the ability to set reminders would revolutionize the lives of many who do not see.

Lists

Lists are a capability currently under investigation for JotChat. We believe this will be a powerful tool, especially for visually impaired users who do not have the same ability to easily make and keep print lists, or who may not be conversant with the numerous computer applications often used to store this information. A generic list capability would enable users to manage grocery lists, general shopping lists, to do lists, things to sell, places to visit, etc. all from the JotChat interface. A secret keeper list could keep lists of passwords, bank accounts, or other information that would require identification before being divulged. Special media lists could intelligently link lists of books, movies, songs, etc to library catalogs, iTunes, etc. One user mentioned keeping lists of books to read as something she stopped doing after losing her sight, but might again be possible to do with a future version of JotChat.

Other comments from the users

In addition to having users rank future JotChat applications, we also asked open-ended questions such as what they liked, what they didn't like, did they like this way of interacting with a computer and how it compares with current ways they keep track of information. Appendix C on page 58 contains a summary of the users' responses.

Testers saw the use of JotChat as a positive experience. They liked the ease of use, being able to "ask it things in the way that you would use everyday speech, instead of figuring out some cryptic word." Users said they would like it as a way to keep them organized. They liked that it could associate things. They also appreciated that "everything can be in one place, one application for everything" eliminating the need to learn 5 applications. Users described it as "fascinating", "smart", "intuitive" and "friendly."

As to what they didn't like, most testers didn't have anything specific about what they disliked about the software. Several mentioned particular things it hadn't understood, which will be included in the next set of sentences to work on. Visually impaired users wanted backspaced characters to be spoken. A few felt uncertain as they figured out the conversational interface, but became easier as the testing progressed. Some suggested additional features like not needing apostrophes, scheduling appointments, or having a thesaurus. Some did not like how it said names. A couple didn't like the specific keyboard and computer synthesized voices used in the test.

Many users asked us when JotChat would be available. Most volunteered to test again or be early adapters. Seeing the user's enthusiasm for this type of product and getting their feedback was valuable and motivating, something that the team will long remember as their work continues.

Appendix A - JOTCHAT TEST SCRIPT FOR ROUND TWO

SETUP FOR THE DAY

Subject packets

- Consent forms (with blank for payment).
- Checks.
- Instruction text for subject to keep.
- Evaluation questions with list of JotChat uses for rating.

Prepare equipment

- Setup test computer with:
- External keyboards.
- VXi TalkPro Headset.
- Speakers and play back device.
- Turn off phones.

SETUP FOR EACH SUBJECT

Before subject arrives if possible

- Fill in payment blanks.
- Transportation/Taxi issues?
- Set up new notes/log files w/subject's test code.

When subject arrives

- Welcome the subject.
- Have subject sign the consent form (Karen will read to VI subjects).
- Play the introduction to subject (Hard copy in subject packet).
- Have VI subjects listen to the computer speech to make sure it is at the right speed and volume.

PRACTICE TESTS OF STATE CAPITALS.

Please type, "What is the capital of Iowa." Note that this type of "capital" is spelled with an "al" at the end. Press the enter key when you finish typing the sentence. Right: the answer is Des Moines.

Now ask JotChat for the capital of any state you wish and press the Enter key. Once again, JotChat will give you the answer.

SCRIPT FOR UNKNOWN WORD DIALOGS

(This text will not be read at this point. It will be used only if/when an unknown word dialog appears during the test session).

(Note to tester. If 1 or both of the likely dialog boxes pop up, we will help the subject work through them on the first 2 occasions they occur. After this, we will simply have them ignore the dialogs and help them get back to the input field).

Because JotChat needs to understand what you say, much as a person would do, it has to understand each word you use. At this time, it does not recognize the word <insert the word they typed>. Thus, you have 3 choices of things you can do. If the word is misspelled, you can ask JotChat for suggestions (the shortcut key for this is Alt S), If the word is the name of a person or place, you can add it to JotChat with Alt A, or you can ignore the word by pressing the Escape key and continue with this scenario. If you do ignore the word, you can either edit the text you have written to change the word, or you can press F3 to delete the text and start over. (or F6 to delete the word)

KEYBOARD ONLY SCENARIOS

(Neal reads a short intro followed by the 17 keyboard only scenarios.)

OK, I think we are ready to begin. Just remember that nothing you do will be judged. We don't care how slowly you type, how many mistakes you make, or whether or not JotChat can't find the answer to your question. We are testing JotChat's ability to correctly respond to your questions. There is a lot it doesn't yet know, so your work today will help us make it smarter.

Do you have any questions at this time?

OK, I will read some scenarios about things I want you to ask JotChat. I won't tell you exactly what to type because we want to discover the many different ways people use to ask questions and retrieve information. Remember to use simple, everyday language and try to type complete sentences.

Ready? Let's begin.

Recycled Scenarios

1. You need to call your friend Paul but you don't know his phone number. What would you type to get this information from JotChat?
2. JotChat knows that Paul has a wife. How do you find who it is?
3. What if JotChat does not have the phone number of your friend, Alice. It is 221-4545. How would you enter this information?
4. Verify that JotChat now has Alice's phone number.

5. Bob works at a computer store. How would you ask JotChat for his number in order to contact him at work?
6. What if you wanted to know all of Bob's phone numbers? How would you get this information?
7. Your friend Paul's cell phone number is 222-3333. How would you give JotChat this information?
8. Bob's email address is bob@nomail.com. How would you enter this in JotChat?
9. You know Paul has a nickname but you can't remember it. Can you find this out from JotChat?
10. A while back you told JotChat about Jim, but now you can't remember who he is, how would you have JotChat jog your memory?

New Scenarios

11. How would you get Mary's address from JotChat?
12. You can also give JotChat addresses, but you need to put a quote at the beginning and end of the address. Also, JotChat will not yet recognize abbreviations, so completely spell out everything in the address. Given that, how would you enter Larry's address, which is:
111 Main Street, Madison, Wisconsin 53700
13. How would you ask JotChat to come up with names of people who live in Madison?
14. How would you ask JotChat for the company that Bob works for?
15. You'd like JotChat to give you a list of all the people you've entered who work at Cool Toys. What would you ask?

List Scenarios

16. JotChat will be able to keep a list of things you need to do or get. If you wanted to have an item, say you are out of milk, appear on such a list, what would you tell JotChat?
17. How would you have JotChat display the list?

SPEECH INPUT ONLY: KEYBOARD EVALUATION QUESTIONS

For subjects doing speech input we will ask the questions 1-4 from the evaluation form before proceeding to the speech input part of the test. This is so that their answers are not biased by their speech input experience. For subjects doing keyboard only testing, we will ask all the evaluation questions after the last scenario.

SPEECH INPUT ONLY: TRANSITION TO SPEECH INPUT

(Neal will start the transition while Karen sets up a new user in Dragon.

- Use new user defaults except training = none
- Make sure microphone volume is max and microphone is on!

When the recording finishes, Karen will assist the user in putting on the headset.)

(This will be a recorded file)

Now, we would like to have you try some additional scenarios using speech input. Instead of typing them, you will be speaking them into a microphone. They will appear in JotChat just like they did when you typed them. There are a few things you need to know about using speech to talk to JotChat.

1. We will be giving you a headset with a microphone attached. We will help you get it placed correctly so that it is optimized for speech. You will not hear anything from the head phone. Rather, you will continue to hear the output from JotChat through the speakers.
2. Speech input is not fool proof. So, sometimes you may say something and the computer will not pick it up correctly. This is not your problem. The problem here is that speech input is far from perfect at this point. We are using it to determine just how well it works. So, don't feel frustrated if some of what you say is not transmitted to JotChat correctly.
3. The software we are using to transmit speech to JotChat is called Dragon Naturally Speaking. It is a common dictation application that you may be familiar with. Once we have the headset adjusted we will have you speak two short passages so that Dragon can adjust to your speech.

Please speak normally. Keep the same volume you normally use. While speaking distinctly, do not over-pronounce words. We want you to sound as natural as you can.

SPEECH INPUT ONLY: SPEECH VOLUME AND QUALITY TESTS

Before we continue with Dragon's volume and quality tests we should make sure:

- The microphone is placed correctly.
- We explain how the mute button works.
- We explain to the visually impaired subjects that they have two options for text to read for these tests.
 1. Read the Braille we had done for the training screens.
 2. Speak anything they wish extemporaneously – counting, reciting the alphabet, etc.

Walk subject through Dragon's speech volume and quality tests. Capture speech to noise ratios

SPEECH INPUT ONLY: ASSESS NEED / DO SPEECH TRAINING

(Have the subject begin dictating the following 10 sentences into Notepad. For visually impaired we will do whisper prompting. We will have a printed copy for sighted users to read)

(If and when a subject dictates 3 sentences in a row that Dragon is able to transcribe perfectly, we will go on to the speech scenarios without further training. Script to explain this to subjects follows)

Dragon recognizes some people's voice quite quickly. Other voices need to be trained. Again, this has nothing to do with how pleasant your voice is. It's just the nature of Dragon. We need to see if your voice needs to be trained. To do this, we will have you read part or all of 10 sentences WE will read aloud to you. When you are done with this exercise, we will either help you with the training or you will go directly to entering the scenarios into JotChat.

Speech Assessment Sentences

1. Larry answers the telephone at work.
2. Paul prefers well behaved children.
3. Whose birthday party is Kelly going to?
4. Paul's mother has an out of town address.
5. Its cool to listen to loons.
6. I can never remember my zip code.
7. When Mary was little she had lots of toys.
8. Alice went down the rabbit hole.
9. Do not use cell phones while driving.
10. Bob does not like reading email.

(If speech training is indicated, Karen will walk subject through the training. Visually impaired subjects can use the Braille printouts or do whisper prompting. Notes should capture whether training was done.)

SPEECH INPUT SCENARIOS

(Subjects not using speech input will continue from scenario 17 to 18 below without interruption, and will use the keyboard for all the scenarios.)

(Subjects using speech input should be told they will now go back to responding to scenarios in JotChat using speech. Tips for speech could be reviewed as appropriate. Other considerations in using speech with JotChat:

- If people who did not train have problems with speech during the test, we may decide to train them. (play this by ear).

- We will start by pressing return for the subject when they finish dictating the sentence. This simplifies what the subject must do physically and allows us to inspect what Dragon transcribes before it is submitted to JotChat. If a subject master this, we may ask that they say "New line" at the end of their sentence.
- Verify that what JotChat receives is what the subject said. If not, the tester should type DErr in front of the sentence and submit it to JotChat so it is put in the log.)

Tell subjects that if Dragon does not accurately transcribe their sentence, we will have them retry three times and then have them use the keyboard.

Speech Input Scenarios

18. How would you ask JotChat for Larry's zip code?
19. What would you ask to get the address of Cool Toys?
20. How would you find out the number of children Kelly has?
21. How would you find out their names?
22. You have never met Bob's mother, but you need to call her. How would you get help from JotChat on this?
23. What would you ask to get Kelly's email address from JotChat?
24. If JotChat could place a phone call for you, how would you ask it to connect you with Bob?
25. Paul is having a birthday soon. Get the date from JotChat.

EVALUATION QUESTIONS

Evaluation questions will be printed on a separate page that includes the list of uses of Jotchat. Subjects that use speech input will be asked 4 questions (5-8) that keyboard only testers are not asked. The two groups are also asked questions at somewhat different times during the test as indicated below.

Subjects doing speech input we will have already been asked questions 1-4 before starting the speech input part of the test. At the end of the test they will be asked questions 5-11.

Keyboard only subjects will be asked their evaluation questions at the end of the test. This includes questions 1-4 and 9-11.

CONCLUSION OF SESSION

- Let subjects know how important their help has been.
- Pay subjects.
- Give visually impaired people money for their cab home.
- Copy JotChat log and back it up on Kathy's computer.

Evaluation Questions for JotChat Usability Tests – Round 2

ALL TESTERS FOLLOWING KEYBOARD TEST

1. What did you like about JotChat?
2. What did you dislike about JotChat?
3. We are going to read you a list of things JotChat may be able to do in the future. Rank how like you would be to use JotChat as a conversational interface for doing the listed tasks

0 – Would definitely NOT use

1 – MIGHT use

2 – Definitely WOULD use

Contacts (4)

	Enter and retrieve names addresses, phone numbers and personal information
	Print address labels
	Get address and other information for stores, restaurants, theaters, etc via MapQuest, on-line phone books or similar services
	Family tree keeper

List management (4)

	Create and maintain shopping and to do lists
	Secret keeper (credit card numbers, social security numbers, passwords)
	Maintain lists of books and other media including interface to library catalogs, iTunes, etc.
	Recipes

Calendar Functions (5)

	Enter and retrieve date and times of appointments, birthdays, due bills, etc
	Set up reminders about appointments, birthdays, due bills, etc
	Synch with Microsoft Outlook
	TV and radio listings
	Bus schedules

Miscellaneous (5)

	Conversational weather queries and reports
	Conversational interface to bring up new or previously identified web pages (Go to Trace's website)
	Conversational interface to general web information sources such as dictionaries, encyclopedias or wikipedias (Read me the wikipedia entry that talks about the origin of philosophy)
	Appliance front end
	Bar code reader for grocery stores

4. Is there anything else you'd like to use JotChat for that wasn't on the list?

SPEECH INPUT ONLY

5. What did you like about using speech with JotChat?
6. What did you dislike about using speech with JotChat?
7. Do you prefer keyboard input or speech input for these types of questions?
8. If JotChat were able to use speech input, how would that affect the tasks you might use it for?

ALL TESTERS AT END OF TEST

9. Did you like this way of interacting with the computer? How does it compare with the current ways you keep track of this type information?
10. Do you have any additional questions or comments?
11. Would you be interested in receiving additional information on this project?

Appendix B - USERS RESPONSES BY SCENARIO**Scenario 1**

You need to call your friend Paul but you don't know his phone number. What would you type to get this information from JotChat?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
7	What is Paul's phone number?
3	What is Pauls phone number?
2	What is Paul's number?
1	Give me Paul's phone number.
1	Look up Pauls phone number.
1	Look up Paul's phone number.
1	Pauls number.
1	Pauls phone number.
1	Paul's phone number.
1	What is the phone number of Paul?
1	What's Paul's number?
11	Unique understood
20	Total understood

Non-understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
0	Unique non-understood
0	Total non-understood

Scenario 2**JotChat knows that Paul has a wife. How do you find who it is?**

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
7	Who is Paul's wife?
4	What is Paul's wife's name?
2	Pauls wifes name.
2	Who is Pauls wife?
1	Pauls spouse.
1	Paul's wife.
1	Pauls wife's name.
1	What is Pauls wifes name?
1	What is Paul's wifes name?
1	What is the name of Pauls wife?
1	What is the name of Paul's wife?
1	Who is his wife?
12	Unique understood
23	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	Paul's wife is?
1	Please look up Pauls wifes name.
1	Whats Pauls wife name?
1	What's Paul's wife's name?
1	Who married Paul?
1	Who's Paul's wife?
6	Unique non-understood
6	Total non-understood

Scenario 3

What if JotChat does not have the phone number of your friend, Alice. It is 221-4545. How would you enter this information?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
8	Alice's phone number is 221-4545.
6	Alices phone number is 221-4545.
3	Alice's number is 221-4545.
2	The phone number for Alice is 221-4545.
1	Phone number for Alice is 221-4545.
5	Unique understood
20	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	Alice number is 221-4545.
1	Alice number:254-5643.
1	Alice phone number 221-4545.
1	Alice phone number is 221-4545.
1	Alices new phone number.
1	Enter Alices phone number 221-4545.
1	Enter Alices phone number.
1	Enter the phone number for Alice.
1	I need to enter Alice's phone number.
9	Unique non-understood
9	Total non-understood

Notes on phone numbers: Phone numbers are separated by dashes, spaces, and sometimes, not separated at all.

Scenario 4**Verify that JotChat now has Alice's phone number.**

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
10	What is Alice's phone number?
3	What is Alice's number?
3	What is Alices phone number?
1	Alices number.
1	Alice's number.
1	Alices phone number.
1	Alice's phone number.
1	Give me Alice's phone number.
1	Give me the phone number for Alice.
1	Please give me Alices phone number.
10	Unique understood
23	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	Alice phone number.
1	Alice.
1	Alice's phone.
1	What is Alice phone number?
1	What is Alice's phone?
5	Unique non-understood
5	Total non-understood

Scenario 5

Bob works at a computer store. How would you ask JotChat for his number in order to contact him at work?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
6	What is Bob's work number?
3	What is Bob's phone number at work?
3	What is Bobs work phone number?
2	What is Bobs work number?
2	What is Bob's work phone number?
1	Give me Bobs phone number at work.
1	Please give me Bob's work phone number.
1	What is Bob's number at the computer store?
1	What is Bobs phone number at work?
1	What is Bob's work phone number at the computer store?
10	Unique understood
21	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
2	Bobs work number.
2	Bobs work phone number.
1	Bobs phone number at work.
1	Bob's work phone number.
1	What Bobs work number?
5	Unique non-understood
7	Total non-understood

Scenario 6**What if you wanted to know all of Bob's phone numbers? How would you get this information?**

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
4	What are Bob's phone numbers?
3	What are all of Bob's phone numbers?
1	All of Bob's phone numbers.
1	Bob's number?
1	Give me all of Bobs numbers.
1	Give me all of Bobs phone numbers.
1	Give me all of Bob's phone numbers.
1	List all of Bobs phone numbers.
1	List Bob's numbers.
1	List Bobs phone numbers.
1	List Bob's phone numbers.
1	What are all Bob's numbers?
1	What are all of Bobs phone numbers?
1	What are Bobs phone numbers?
14	Unique understood
19	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	All phone numbers for Bob.
1	Bob's number are?
1	Bobs phone all.
1	Can I have all of Bob's phone numbers?
1	List all the phone numbers for Bob.
1	List the home and work numbers for Bob.
1	What is all Bob's information?
1	What is Bob's contact info?
1	What is Bob's contact?
1	What phone numbers will call Bob?
1	What phone numbers will reach Bob?
11	Unique non-understood
11	Total non-understood

Scenario 7

Your friend Paul's cell phone number is 222-3333. How would you give JotChat this information?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
8	Pauls cell phone number is 222-3333.
7	Paul's cell phone number is 222-3333.
2	Paul's cell phone is 222-3333.
2	The cell phone number for Paul is 222-3333.
1	Pauls cell number is 222-3333.
1	Paul's cell number is 222-3333.
6	Unique understood
21	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	Cell phone number for Paul 222-3333.
1	Paul cell 222-3333.
1	Pauls cell is 222-3333.
1	Paul's cell phone number 222-3333.
4	Unique non-understood
4	Total non-understood

Scenario 8**Bob's email address is bob@nomail.com. How would you enter this in JotChat?**

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
7	Bobs email address is bob@nomail.com.
4	Bob's email address is bob@nomail.com.
3	Bob's email is bob@nomail.com.
2	Bob's e-mail address is bob@nomail.com.
1	bob@nomail.com is Bob's email address.
1	Bobs e-mail address is bob@nomail.com.
1	Bob's email address is: bob@nomail.com.
1	Bob's e-mail is bob@nomail.com.
8	Unique understood
20	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	Bob e-mail is bob@nomail.com.
1	bob@nomail is Bob's email address.
1	bob@nomail.com is the email address for Bob.
1	Bob's email bob@nomail.com.
1	The email address for Bob is bob@nomail.com.
5	Unique non-understood
5	Total non-understood

Scenario 9

You know Paul has a nickname but you can't remember it. Can you find this out from JotChat?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
9	What is Paul's nickname?
7	What is Pauls nickname?
1	Give me Paul's names.
1	Look up Paul's nickname.
1	Pauls nickname.
1	What is Paul's nick name?
1	What is the nickname of Paul.
7	Unique understood
21	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	Does Paul have a nickname?
1	What is the nickname for Paul?
2	Unique non-understood
2	Total non-understood

Scenario 10a

There is someone named Jim, you want to find some information about him. How would you begin?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
13	Who is Jim?
3	Tell me about Jim?
2	Who is Jim Rockford?
1	Tell me about Jim Eastman.
1	Tell me all about Jim.
1	What is Jim Rockford's phone number?
1	Where does Jim live?
1	Where does Jim work?
8	Unique understood
23	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	Can you give me information about Jim?
1	Find Jim.
1	Give me Jim Eastman's address and phone numbers.
1	How old is Jim Rockford?
1	Information about Jim?
1	Jim?
1	Jim's information?
1	Please give me Jim Eastman's information.
1	Please give me Jim's information.
1	Tell me about my acquaintance Jim.
1	Tell me about my friend Jim.
1	What can you tell me about Jim Rockford?
1	What do you know about Jim?
1	What information do you have about Jim?
1	What Jim Eastman's information?
15	Unique non-understood
15	Total non-understood

Scenario 10b**Which Jim?****1. Jim Eastman ;****2. Jim Rockford ;**

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
7	1
6	2
3	Jim Eastman.
2	Tell me about Jim Eastman.
1	Jim Rockford.
1	Who is Jim Rockford?
6	Unique understood
20	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	Find Jim Rockford.
1	Give information about Jim Eastman.
1	Give me information on Jim 1.
1	Information on the first Jim.
1	Jim that I met.
1	Tell me about both
1	Two.
1	What do you know about Jim Eastman.
8	Unique non-understood
8	Total non-understood

Scenario 11**How would you get Mary's address from JotChat?**

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
6	What is Marys address?
6	What is Mary's address?
2	Please give me Mary's address.
2	Where does Mary live?
1	Give me Mary's address?
1	Look up Mary's address.
1	Marys address.
1	Mary's address.
1	What is Marys street address?
9	Unique understood
21	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	May I have Mary's address?
1	What is Marry's address?
1	What is Marys complete address?
1	What is the address for Mary?
4	Unique non-understood
4	Total non-understood

Scenario 12

You can also give JotChat addresses, but you need to put a quote at the beginning and end of the address. Also, JotChat will not yet recognize abbreviations, so completely spell out everything in the address. Given that, how would you enter Larry's address, which is: 111 Main Street, Madison, Wisconsin 53700

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
7	Larry's address is "111 Main Street Madison Wisconsin 53700".
2	Larry's address is "111 Main Street Madison Wisconsin 53700.
4	Larrys address is "111 Main Street Madison Wisconsin 53700".
1	Larrys address is "111 Main Street Madison Wisconsin 53700.
1	Larrys address is "111 Main Street Madison Wiisconsin 53700.
1	Larrys address is "111 Maine Street Madison Wisconsin 53700".
1	Larrys address is "111 Main Street Madison Wisconsin 53770".
1	Larrys address is "111 Main Street Madison, WI, 53707".
1	Larry lives at "111 Main Street, Madison, Wisconsin 53700".
1	Larry's home address is "111 Main Street, Madison, Wisconsin 53700".
10	Unique understood
20	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	"Larrys address is 111 Maine St. Madison, Wisconsin 53700".
1	Enter Larry's address "111 Main Street Madison Wisconsin 53700.
1	Larry's address "111 Main Street Madison Wisconsin 53700.
3	Unique non-understood
3	Total non-understood

Scenario 13**How would you ask JotChat to come up with names of people who live in Madison?**

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
16	Who lives in Madison?
1	My friends who live in Madison.
1	People who live in Madison.
3	Unique understood
18	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
3	Who lives in Madison Wisconsin?
1	Does Mary live in Madison?
1	Give me all names of people in Madison.
1	Give me all of my friends who live in Madison.
1	Give me name of people in Madison.
1	Give me names of people who live in Madison Wisconsin.
1	Give me names of people who live in Madison.
1	Give me names of three people who live in Madison.
1	Give me the names of people living in "Madison, Wisconsin".
1	List people in Madison, Wisconsin.
1	List the people you know of who live in Madison.
1	Make a list of Madison residents.
1	Names of all people who live in Madison.
1	People who live in Madison Wisconsin.
1	Please give me names of individuals living in "Madison, Wisconsin".
1	Please give me names of people living in "Madison, Wisconsin".
1	Please give me the names of people living in "Madison, Wisconsin".
1	What are the addresses of people who live in Madison?
1	What are the names of people living in "Madison, Wisconsin"?
1	What are the names of people who live in Madison, Wisconsin?
1	What are the names of people who live in Madison?
1	What names have I given for people who live in Madison?
1	Where is Larry?
1	Who are the people who live in Madison?
1	Who do I know that lives in Madison?
1	Who lives in Madison, Wisconsin?
26	Unique non-understood
28	Total non-understood

Scenario 14**How would you ask JotChat for the company that Bob works for?**

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
11	Where does Bob work?
5	What company does Bob work for?
3	Bob works for what company?
3	Who does Bob work for?
4	Unique understood
22	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
2	What is the name of the company Bob works for?
1	Bobs companys name.
1	Company Bob works for.
1	Name of the company Bob works at.
1	Name of the company where Bob works.
1	What is Bobs companies name?
1	What is the company Bob work for?
1	What is the company Bob works for?
1	What is the name of the company that Bob works at?
1	Which company does Bob work for?
10	Unique non-understood
11	Total non-understood

Scenario 15

You'd like JotChat to give you a list of all the people you've entered who work at Cool Toys. What would you ask?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
19	Who works at Cool Toys?
1	Who works for Cool Toys?
2	Unique understood
20	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	Cool Toys employees.
1	Cool Toys workers.
1	List all people that work at Cool Toys.
1	List employees at Cool Toys.
1	List people at Cool Toys.
1	List the people who work at Cool Toys.
1	List who works at Cool Toys.
1	Name all people that work at Cool Toys.
1	Names of people that work at Cool Toys.
1	Please give me a list of all employees of Cool Toys.
1	Please give me a list of people who work at Cool Toys.
1	Who all works at Cool Toys?
1	Who else works at Cool Toys?
13	Unique non-understood
13	Total non-understood

Scenario 16

JotChat will be able to keep a list of things you need to do or get. If you wanted to have an item, say you are out of milk, appear on such a list, what would you tell JotChat?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
4	Put milk on grocery list.
3	Put milk on my list.
2	I need milk.
2	Put milk on the grocery list.
1	I need more milk?
1	Milk is on the shopping list.
1	Please put milk on my grocery list.
1	Put "milk" on grocery list.
1	Put milk on list.
1	Put milk on the list.
1	Put milk on the shopping list.
11	Unique understood
18	Total understood

Non-understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
3	Add milk to grocery list.
3	Add milk to list.
2	Add milk to the list.
2	Grocery list milk.
2	I need to buy milk.
2	We need milk.
1	Add "milk" to grocery list.
1	Add "milk" to my list.
1	Add milk to groceries.
1	Add milk to my grocery list.
1	Add milk to my shopping list.
1	Add milk to shopping list.
1	Add milk to the grocery list.
1	Add milk to the shopping list.
1	Add milk to weekly grocery list.
1	Add the item milk to the list.
1	Go shopping for milk and cat food.
1	Grocery list should include milk.
1	I need to go to the store to buy milk.
1	I want milk from the grocery store.
1	I want to add milk to my shopping list.

1	Include milk in my grocery list.
1	Include milk in the shopping list.
1	Kopps grocery list milk.
1	List should include milk.
1	Milk.
1	Please add milk to my list.
1	Remember to buy kitty food and tooth paste.
1	Remember to get cat litter.
1	Shopping list: milk.
1	Woodmans grocery list milk.
31	Unique non-understood
39	Total non-understood

Scenario 17**How would you have JotChat display the list?**


Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
5	What is on the grocery list?
4	What is on my grocery list?
3	What do I need?
3	What is on my list?
2	What is on the shopping list?
1	What is on grocery list?
1	What is on list?
1	What is on the list?
8	Unique understood
20	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
4	Show grocery list.
3	Grocery list.
2	List grocery list.
1	Display grocery list.
1	Display list.
1	Display my list.
1	Do I need milk?
1	Give me grocery list.
1	Is milk in the database?
1	List groceries.
1	List items on grocery list.
1	List needed groceries.
1	List.
1	Print shopping list.
1	Remember my grocery list.
1	Say list.
1	See grocery list.
1	Show list.
1	Show me grocery list.
1	Show me my grocery list?
1	Show my list.
1	Show shopping list.
1	Tell me what is on my grocery list.
1	The shopping list is what?
1	View grocery list.





1	What do i need to buy at the store?
1	What do we need?
1	What is listed on the grocery list?
1	What is my list?
1	What is the grocery list?
1	Where is milk?
31	Unique non-understood
37	Total non-understood

Scenarios 18 – 25 were input via speech by 13 of the 20 users. The remainder of the users continued to use keyboard to enter their responses.  indicates the response was given by a keyboard user, a superscript indicates more than one keyboard user entered this response.


Scenario 18

How would you ask JotChat for Larry’s zip code?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
13	What is Larry's zip code?  2
3	What is Larrys zip code?  3
1	Larrys zip code.
1	What is Larry's zipcode?
1	What is the Larry's zip code?
1	What's Larry's zip code? 
1	Where does Larry live? 
7	Unique understood
21	Total understood






Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	What is Larry zip code?
1	What is Larry's zip? 
2	Unique non-understood
2	Total non-understood





Scenario 19

What would you ask to get the address of Cool Toys?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
8	What is Cool Toys' address?  2
8	What is the address of Cool Toys?  3
5	What is the address for Cool Toys? 
1	Cool Toys' address is what?
1	Give me the address for Cool Toys. 
1	Tell me about Cool Toys.
1	Where does Cool Toys live? 
7	Unique understood
25	Total understood




Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
3	Where is Cool Toys?  2
1	Cool Toys address. 
1	Cool Toys is where?
1	What is Cool Toys address? 
1	What is the address Cool Toys?
1	What is the address where Bob works? 
6	Unique non-understood
8	Total non-understood



Scenario 20

How would you find out the number of children Kelly has?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
15	How many children does Kelly have?  4
3	How many kids does Kelly have?  2
1	Kelly has how many children?
1	List Kelly's kids. 
4	Unique understood
20	Total understood






Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
1	Have many children does Kelly have? 
1	How many children belong to Kelly? 
2	Unique non-understood
2	Total non-understood





Scenario 21

How would you find out their names?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
8	What are the names of Kelly's children?
2	What are Kelly's children's names? 
2	Who are Kelly's kids?  2
1	Names of Kellys children. 
1	The names of Kelly's children are what?
1	What are her children's names?
1	What are Kellys childrens names? 
1	What are Kelly's kids' names?
1	What are the name of Kelly's children?
1	What are the names of Kellys children? 
1	Who are Kelly's children?
11	Unique understood
20	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
3	What are their names? 
2	What are Kelly's kids names? 
1	What are Kellys kids names? 
1	What are Kelly's kid's names? 
4	Unique non-understood
7	Total non-understood

Scenario 22

You have never met Bob's mother, but you need to call her. How would you get help from JotChat on this?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
5	What is Bob's mother's phone number? 🗨️
2	What is Bob's mom's number? 🗨️ 2
2	What is Bob's mom's phone number? 🗨️
2	What is Bob's mother's telephone number?
2	What is the phone number for Bob's mother?
1	Give me Bobs mother's phone number. 🗨️
1	Tell me Bob's mom's phone number?
1	What is Bob's mother's number?
1	What is Bobs mother's phone number? 🗨️
1	What is the name of Bob's mother? What is Emily's phone number?
1	What is the phone number of Bob's mother?
1	Who is Bob's mom? 🗨️
1	Who is Bob's mother? What is Emily's phone number? 🗨️
13	Unique understood
21	Total understood






Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
2	What is Bobs mothers phone number? 🗨️ 2
1	Bobs mother has a phone number What is it. 🗨️
1	Bobs mother has a phone number. What is it? 🗨️
1	Bobs mothers phone number. 🗨️
1	Give me Bobs mothers phone number. 🗨️
1	What is Bob's mother's name and phone number?
1	Who is Bob's mom and what is Bob's mom's phone number? 🗨️
7	Unique non-understood
8	Total non-understood


Scenario 23

What would you ask to get Kelly’s email address from JotChat?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
13	What is Kelly's e-mail address?  2
2	Give me Kellys email address.  2
2	What is Kelly's email address?  2
1	Give me Kelly's email address. 
1	Give me Kelly's e-mail address.
1	What is Kelly's e-mail adress? 
6	Unique understood
20	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
3	What is Kelly's e-mail?  2
1	Kelly's e-mail address.
1	Kelly's e-mail.
3	Unique non-understood
5	Total non-understood

Scenario 24

If JotChat could place a phone call for you, how would you ask it to connect you with Bob?

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)







Count	Response
8	Call Bob. 📞 4
6	Please call Bob. 📞
2	Call Bob at work. 📞
1	Call Bob on the phone.
1	Call Bob's telephone number.
1	Connect me to Bobs phone number. 📞
1	Please dial Bob for me.
7	Unique understood
20	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
0	Unique non-understood
0	Total non-understood

Scenario 25**Paul is having a birthday soon. Get the date from JotChat.**

Understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
12	When is Paul's birthday? 
6	What is Paul's birthday?  3
2	When is Pauls birthday?  2
2	When was Paul born? 
1	What is Paul's birthdate? 
1	What is Pauls birthday? 
6	Unique understood
24	Total understood

Non- understood user inputs (includes any variation of capitalization, commas or end of sentence punctuation)

Count	Response
2	What is the date of Paul's birthday?
1	When is Paul's bday?
2	Unique non-understood
3	Total non-understood

Appendix C – WHAT THE USERS LIKED, DISLIKED AND INFORMATION THEY WANT IN JOTCHAT

Things users said they liked:

It was fun. I like that it could keep me really organized. I like trying out new equipment too.

It's just fascinating. It's pretty smart.

I like the idea that you can ask it things in the way that you would use everyday speech, instead of figuring out some cryptic word. For somebody who isn't used to using a computer, this would help a lot. You don't have to be uncomfortable with it.

It was giving me answers.

I thought that it was more intuitive than before. I could ask it more directly, without verbs, which is good.

Liked that it was pretty simple wording. If you are used to using Google, you have to leave out words. I like shortcuts to be there, but my grandparents would like the longer explanations. It was nice that I could just ask for things. Tester also liked that it didn't depend on syntax.

It is a more sophisticated thing than what I went through this morning with Verizon, an automated call attendant. It was fairly clever so far, a clever idea.

Likes that you don't have to be verbose to get an answer. You can be "short & sweet." It can associate things. It is very easy to use. You don't have to learn 5 applications. It is nice that you don't have to put in caps. Likes that everything can be in one place, on application for everything.

Liked the times when it said there were two whatever, handling lists and groups. It doesn't clutter them up with more numbers.

The thing that I found most interesting is the underlying conceptual stuff about how language is built, how understanding happens. Tester says that if it was available, they would use it.

It can use regular words when asking questions. It gives you information without having to look at a whole chart, a list, or something you have to read.

The speech is clear. The scenarios seem clear and it seems quite able to do that.

Tester likes the idea of a grocery list. Likes that they don't have to navigate to anything in JotChat. Don't have to remember where it is, just ask for it.

Finds the concept very interesting and challenging to understand its modus operandi, how it structures logic.

I really like that you are talking to it in English. I really like that. It's great, a great application. It's friendly. It would be great, especially if it was speech controlled.

Tester thought that it was easy and liked that you could type in and it would speak to you. When you are visually impaired, you have to look for it.

It is pretty quick. You ask a pretty natural question, and as a user you quickly learns its bounds, no Boolean. It's case insensitive. It gives you feedback on what it doesn't know.

Tester liked that they could get information and put it in. Tester likes using the keyboard. Tester liked showing possession.

I think that it is kind of neat, that you can ask phone numbers. Tester thinks that the speech is good. I like how you can ask it questions.

Like that it is easy to use and that it seems like in the future it will be fairly intuitive. I like that you can talk to it in every day speech. It's not formal. You can just type in so-and-so's address.

Things users said they disliked:

Nothing. I just didn't like this specific keyboard.

There's nothing. I don't dislike anything. It just fascinates me. The parameters you set up, the rules. There are several ways to ask a question.

Trying to figure out what it is going to figure out and what it isn't. It's a process, that comes with time.

Can store everything tester wants on cell phone. But doesn't dislike anything. Thinks that it is a really good idea.

I don't think that there was anything that I disliked. I was trying to shortcut right away.

When it told me that it didn't understand a word and wouldn't let me keep typing, when it stops me in the middle. Tester wishes that the list function worked better because the tester would use that a lot.

It should have a thesaurus, a registry, for words, so that it would understand alternative words, like "wife", "mother", "aunt" etc.

It would be nice if you didn't have to put in apostrophe.

Tester can't tell when they've made a mistake or what kind of mistake they made. We didn't set up their session for echo mode, but that would still leave us with the "why" question.

Tester didn't like the keyboard used. Tester didn't like the speech voices because they weren't used to them. Tester is used to the JAWS voice.

It didn't seem to quite understand me sometimes. Tester is working through the idea of using fewer key strokes with part of the tester's mind shifting to speech input.

I was expecting an appointment question. It would be interesting to see how that would work. It would be nice to have that in a natural language. Tester would like both speech and keyboard entry.

Nothing except the way that it brought up last names first and then the first names. Specificity. It was very picky, specific, to retrieve input in some areas, not generally. That it didn't know how to add something to a list. I see it as a prototype, so I really don't have a dislike. It will be great. It already seems like it would be useful, definitely.

I didn't really dislike anything. It was just learning what to say.

Computer voices are computer voices, but this is better than Microsoft Mike. I can't say that there is anything that I don't like.

Tester couldn't figure out its way of thinking. First, the tester thought that they could be really simple, but that didn't always work. Tester felt that they needed to use proper English more.

I don't like when you backspace it removes the full word. Can it backspace with each letter? The echo key should be changed and the tester asked for the speed of the voice to be increased. It seems to be pretty intuitive, but it would be nice to have a help. Tester would like it to understand more, like "please add milk to my grocery list."

The more I heard the voice, the less that I liked it. It eats its words. It is garbled, especially if you go faster. Tester suggests "Heather" from AT&T Mobile. Tester feels that they don't know enough about it to really not like it. Tester wonders if it would be better if it didn't "yell at you" when you misspelled something. Tester would like to see better error handling.