

The User Context: Aligning System and User Behavior

4/30/10 Position Paper:

W3C Workshop on Conversational Applications -- Use Cases and Requirements for New Models of Human Language to Support Mobile Conversational Systems.

Kimberly Patch
Redstart Systems
kim@redstartsystems

Many people in the accessibility community depend on speech input, and many have been using it either in tandem with other input methods or completely hands-free for many years. Many have also been frustrated by it.

There's a wealth of real-world information in this community that can provide valuable insights into future speech standards.

User experience can also change with different types of use: evaluation, intermittent, use for specific applications like search, use in tandem with other input methods, and varying degrees of hands-free access. Just as application contexts matter with speech, so do user contexts.

Here are some things to think about:

- Study: Predictors of Abandonment of Speech Recognition Technology in Children and Adolescents with Physical Disabilities
Tara Previl and Amita Furgoch at the Ottawa Children's Treatment Center interviewed 31 children and adolescents with physical disabilities who need to use and received training in desktop speech recognition software
Results: speech recognition software has a 45% abandonment rate even among this highly motivated community
Abandonment factors included vocal fatigue and frustration

Why is talking to a device more difficult than talking to a human? What can we do to address this?

Why is talking to a device more frustrating than talking to a human? What can we do to address this?

- Speech input is different enough from traditional keyboard/mouse/touch input that speech users encounter unique or exacerbated interface problems. Focus issues, for instance, are trivial for mouse/touch users because moving the arrow automatically changes the focus. Focus causes bigger problems for speech users not only because they tend to use the mouse arrow less often, but also because speech commands can depend more heavily on context. Focus problems that traditional mouse and keyboard users might not notice can be major disruptions for speech users.

- Speech input is different enough from the traditional input methods that feedback needs can be different. For instance, a mouse/touch user moves the cursor, selects text and deletes it in three separate steps with plenty of time in between to see where the cursor has moved to and what has been selected. If a speech user does all three things at once and there is no extra feedback the user may have to reorient to see what has happened.

- A special point of frustration for speech users is that they can anticipate when speech software isn't going to work well but can't do anything about it. Given an imperfect system, giving the user the means to anticipate a problem and avoid it is important. For instance, when users know from

experience that the system is likely to mistake "2" for "two", they are less frustrated when they have a way to specify what they want.

- Another big complaint is users can't remember commands. Can better means of command discovery, adjustment, organization and sharing address this problem? Can standards make it easier for developers to carry this out?

As the W3C works out a more comprehensive language model, it's important that the standards are flexible enough to allow developers to anticipate and address these known issues, and that they don't inadvertently exacerbate any of them.