My Big Fat Main Menu: The Case for Strategically Breaking the Rules

Susan L. Hura, PhD
SpeechUsability
27 October 2006
Contributed for *VUI Visions*

How many items should there be on a VUI menu? Even for basic questions like this, VUI designers are often forced to derive solutions from research conducted in other fields that we translate to the VUI domain. Often the translation is none too obvious. Case in point: George Miller's 50-year-old research on short term memory (1956) is often portrayed as the golden rule for VUI menu size limits. Miller's research focused on participants' ability to recall items presented auditorily—most people could remember between 5 and 9 items, hence the 'magical number 7' title of his famous paper. The limits of auditory short term memory are real, but they are far from the whole story for how to design a VUI menu. I know this because I have observed many very large menus that are usable, efficient, and well-received by users. How do we explain this seeming inconsistency?

Let's start by defining exactly what users do when they interact with a VUI menu. Users listen to a list of options, consider each one, then select an option they hope will allow them to achieve a goal. This listen-evaluate-decide paradigm is very different from the simple recall task in Miller's studies. The primary purpose of a menu is to enable users select among options quickly and easily. Therefore, I suggest that we should instead focus on designing a set of menu options is *descriptive* and *distinct*, rather than worrying about exactly how many options there are.

Descriptive menu options give users the immediate sense of 'I know what that is!' Recognizable terminology in menu option names enables users to evaluate and discard non-desired options without holding them in memory. Similarly, users will immediately know the difference between *distinct* menu options, which eliminates the need to remember the entire list while making their menu selection. Current short term memory theorists agree that there are many factors beyond the number of items that affect auditory short term memory, such as word length, frequency and familiarity, and inter-item discriminability (that is, how distinct items on a list are from one another; see Nairne, 2002). The importance of descriptiveness is supported by recent research in human computer interaction showing that meaningful menu option labeling is more important than information architecture for users (Resnick & Sanchez, 2004)

Distinctness and discriminability were even more important than usual when designing a recent VUI menu. I started this project with the luxury of an extended period of requirements gathering that included user interviews and analysis of months of call flow statistics. At the end of my analyses, I was left with a shocking conclusion: I proposed a main menu with 34 options. While I'm not a believer in the magical number 7 theory of menu design, 34 seemed ridiculously high. The reasoning that lead to this menu highlights how you can break the rules of VUI design without betraying the overall goal of creating a usable automated system.

Here are the basic facts of the project.

- Project is to design an internal facing IVR for a large US corporation whose purpose is to route employee calls to the appropriate live representative.

- Two distinct groups of users. Field users call multiple times a day. They call to manage every aspect of their job and call almost exclusively from their cell phones. Home office users call once or twice a year. They typically call from their desks after failing to find an answer on the company intranet. They are calling to resolve benefits or human resource issues.

- Each user group currently served by its own IVR systems with largely overlapping functionality. Current IVRs use very large grammars to support an open ended initial question ("*To better assist you, please say a word or phrase for the topic you need.*") Many users are frustrated by not knowing what to say, or because of poor recognition performance, and some users employ techniques to fool the system into taking them to a live agent right away. However, perhaps 25% of calls are users successfully and happily using keyword system.

- The goal of my redesign project is to consolidate the two current systems into one IVR while improving correct call routing. All current functionality (including keywords) must remain available.

- The client rejected building a statistical language model to better support their open-ended initial question.

I came to a few conclusions immediately. First, I decided we need a menu to serve infrequent callers. These users don't know what to say and they are uncomfortable. This was a troubling decision because I knew that a hierarchical menu wouldn't work. Previous usability testing showed that users were unable to choose the intended menu option for many tasks when presented with a top level choice of 'benefits,' 'policies,' 'payroll,' and 'other.' The picture became even more complex when call statistics showed that the most frequently used functionality was a mix of very high level ("computer assistance") and very specific options ("family and medical leave act"). These facts all pointed me towards a wide, shallow menu.

Because this menu would necessarily be longer, I realized that there must be a way to bypass the menu for frequent callers and those who successfully use keywords today. These facts support a keyword recognition strategy, but recall that improved recognition performance is another goal. To accomplish this, I proposed using a modified version of the current keyword grammar at the main menu in which we significantly prune number of items, and use weighting based on frequency of use.

A final consequence of using a wide single-layer menu is that we need one submenu for field sales options because not everyone needs to hear them.

The options on my big fat main menu are grouped into categories; each category ends with an "other" option where a subset of keywords appropriate to that category are recognized.

Users are instructed up front that they can barge in with a keyword at any time, and that they can speak menu options at any time. The voice talent speaks quickly here—the entire menu lasts only 61 seconds.

Here is the Main Menu prompt:

> *You can say....Company Directory. Computer Assistance. Say Field Support to hear those options.*
>
> *For health benefits you can say Unified Healthcare, Rogers, Dental, Healthcare Reimbursement Accounts, Express Refills, or Other Health Benefits.*
>
> *You can also say Payroll, Direct Deposit, W2, Base Compensation, or Other Payroll Questions.*
>
> *For financial benefits you can say 401K, Trading ID, Stock Options, or Other Financial Benefits.*
>
> *You can say Human Resources, Peopletools, Employment Verification, Job Posting, Termination, or Other HR Issues.*
>
> *For company policies, you can say Family and Medical Leave Act, Short Term Disability, Flexible Work Options, Vacation, Retirement, or Other Policies.*
>
> *Or, you can say Representative if you know you need a person.*

What we have observed in practice is that frequent users barge in and never hear the menu. Infrequent users who have a term in mind for what they need also barge in without hearing much of the menu. It's only when the user does not know what word to say that they bother listening to the menu. And in that situation, having a menu is an asset, not a punishment. There you have it: thirty four options, quick and simple, the users like it, the client likes it. And I broke a lot of rules.

The downfall of many rules of VUI design is that they are presented as generic prescriptions to be followed in all cases. This makes the rules both too broad and too specific simultaneously. The rules are overly specific because they focus on, for example, the exact number of items in the menu rather than on the more important principle of designing within the limits of auditory short term memory. Rules are overly broad in that they refer to any main menu, with no regard to differences in domain, user groups, frequency of use, or other factors that should influence menu design.

Here, having 34 items on the menu is no burden to users because the options are distinct and discriminable, and users are not forced to listen to the menu if they don't need it. For these users, on this system, the answer to 'how big should a VUI menu be?' is 34 items, no disrespect to Miller's magical number intended.

Miller, G., 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.

Nairne, J. 2002. Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, 53, 53-81.

Resnick, M. & Sanchez, J. 2004. Effects of organizational scheme and labeling on task performance in product-centered and user-centered web sites, *Human Factors*, 46, 104-117.