

Let's try that again ! ...

Interfaces vocales d'hier à
aujourd'hui ... et vers demain ?

« Mes petits-enfants vont regarder ce qu'était un clavier et ils vont rigoler. »
David Barnes, IBM -
in Home Office Computing -
September 1999.

Avertissement:

Les extraits sonores de cette présentation ont
été désactivés pour cause de propriété
intellectuelle. Nous nous en excusons.

Pour un essai de nos produits vous pouvez
composer le 1-866-TRY-INSP.

Merci !

Pour joindre l'auteur: blackntorck@sympatico.ca

Orientation de la présentation

- Quelques définitions et concepts
- En quoi les interfaces vocales (VUI) diffèrent-elles des interfaces graphiques (GUI) ?
 - Au niveau de la conception
 - Au niveau de leur évaluation
- Exemples de bonnes et mauvaises pratiques

Définitions et concepts

- **Reconnaissance Vocale Automatisée**
 - « *Technologie permettant une communication avec une machine informatique par l'utilisation de paroles (phrases ou mots).* »
 - Helander, Landauer & Prabhu, 1997.
 - Technologie permettant à un système de recevoir le son de la voix humaine, de séparer le signal vocal du bruit ambiant et d'accepter le message de cette voix comme input pour contrôler un système.

Définitions et concepts

■ Reconnaissance Continue vs Échantillonnée

- Reconnaît un **vocabulaire restreint** (et présumément acquis) représentant des fonctions prédictibles du système
- **Speaker dependent**
- Basé uniquement sur la reconnaissance phonologique, et non sur l'interprétation grammaticale ou syntactique
- ViaVoice de IBM:
 - « **Undo this** » -> Annule l'action exécutée
 - « **Copy this** » -> Copie le texte sélectionné
 - « **Select <text>** » -> <text> est un mot ou une phrase qui sera sélectionnée.

Définitions et concepts

■ Reconnaissance Continue vs Échantillonnée

- Permet de s'adresser au système de manière **naturelle**, sans utiliser de commandes spécifiquement acquises pour utiliser ce système
- Introduit la notion de dialogue en « langage naturel » (discuté plus loin)
- **Speaker independent**
- Basé sur la compréhension du langage, mais aussi du sens et du contexte (phonologie, grammaire, sémantique)
- Requier beaucoup de ressources d'interprétation

Définitions et concepts

■ Speaker Dependent

- Requiert une phase « d'entraînement »
- Non flexible parce que limité à un seul locuteur
- Reconnaissance meilleure -> suggéré pour les applications critiques (ex.: contrôle de processus)

■ Speaker Independent

- Pas « d'entraînement », plus flexible, mais limité pour l'instant à des vocabulaires encore assez restreints.
- Cependant, les progrès de la reconnaissance nous permettent d'espérer pour bientôt plus de performance de la part des systèmes indépendants à reconnaissance continue.

Définitions et concepts

■ Word Spotting ou Parsing

- Technique employée dans les systèmes à reconnaissance échantillonnée pour offrir une illusion de reconnaissance continue.
- Filtrage de l'énoncé pour n'interpréter que les actions signifiantes à l'application.
- Semble plus naturel à l'utilisateur:
 - « I want the horoscope for Scorpio please. » ou
 - « I'd like the stock quotes for Cisco Systems. »
- Précurseur du Keyword barge-in ...

Définitions et concepts

■ Interface vocale

■ « Où l'interlocuteur et la machine interagissent par la voie du langage parlé. Le langage parlé remplace le terminal visuel, les boutons, les indicateurs, et les manettes des interfaces traditionnelles. La machine présente de l'information à l'interlocuteur par le biais d'enregistrements ou de texte synthétique - elle récite l'information - et accepte et interprète l'entrée d'info par la reconnaissance de la parole. » - Bruce Balantine, 1999.

■ Dialogue entre un utilisateur et un système où le mode privilégié d'interaction est le langage oral.

Exemples d'interfaces vocales



Fidelity
Investments



American
Airlines



Infospace SMS
Alert Manager

Caractère naturel des VUI

- Aucun apprentissage pré-requis
 - « ... aujourd'hui encore, les êtres humains, en majorité, savent parler sans savoir lire. »
 - A. Martinet, Éléments de linguistique générale
- Suit le déterminisme linguistique qui dit qu'on émet une pensée avec des mots, et qu'il est naturel de verbaliser.
 - « C'est dans le mot que nous pensons... C'est le mot qui donne à la pensée son existence la plus haute et la plus vraie. »
 - Hegel, Philosophie de l'esprit

VUI versus GUI

- En quoi les interfaces vocales (VUI) diffèrent-elles des interfaces graphiques (GUI) ?
 - Au niveau de la conception
 - Au niveau de leur évaluation

VUI versus GUI:

- Quelques distinctions ...
 - Processus centré sur le « dialogue »
 - Domaines d'application
 - Tâches et processus types
 - Affectation des modes
 - Contraintes de conception
 - Scénarisation
 - Guides de style
 - « Challenges »

VUI versus GUI: Conception

- Processus centré sur le dialogue
 - Philosophie
 - La conception de l'interaction doit être faite **oralement** et non visuellement.
 - Étapes
 - Analyse
 - Conception
 - Construction
 - Évaluation
 - Maintenance

VUI versus GUI: Conception

■ Processus centré sur le dialogue

■ Conception

1. Identifier les "tokens"
2. Identifier le vocabulaire technique
3. Concevoir les grammaires
4. Concevoir le dialogue (invites, messages)
5. Concevoir le **Call Flow**
6. Concevoir la livraison de l'information (**Output**)
7. Concevoir la personnalité de l'application
8. Exécuter un **Wizard of Oz**
9. Analyser les résultats du **WOZ**
10. Ajuster la conception de l'interface



VUI versus GUI: Conception

■ Processus centré sur le dialogue

■ Construction

1. Identifier les segments à enregistrer dans le dialogue (invites, messages)
2. Identifier les segments à enregistrer dans le contenu à livrer (output)
3. Produire le document de studio
 - Description de la personnalité et de la prosodie
 - Liste des invites
4. Sélectionner la/les « voix »
5. Procéder aux enregistrements en studio
 - Coaching
6. Faire la post-production des segments
 - Découpage, normalisation, mixage des ambiances sonores ...

VUI versus GUI: Conception

- Processus centré sur le dialogue
 - Évaluation
 1. Tests d'utilisabilité
 2. Analyse des résultats
 3. Ajustements à l'interface en conséquence

VUI versus GUI: Conception

- Processus centré sur le dialogue
 - Maintenance
 - Monitoring des Call logs
 - Identifie les grammaires activées et les OOV, et permet l'ajustement des grammaires
 - Identifie les comportements des utilisateurs non obtenus durant les tests d'utilisabilité
 - Surprise, Impatience, Insatisfaction, Erreurs, Recours à l'aide
 - Performance du système
 - Statut de la reconnaissance, Echo cancellation,
 - Conditions identifiées (Too Soft, Too Soon, Too Noisy...)

VUI versus GUI: Conception

■ Domaines d'application

■ De moins en moins limités:

- Service à la clientèle (traitement / acheminement des appels)
- Standards téléphoniques
- Médical / Légal (transcription de dictée)
- Services d'aide à la productivité (email, agenda, conference call, outils de recherche)
- Militaire et sécurité
- Loisir (info-services, alertes, etc.)

VUI versus GUI: Conception

■ Tâches et processus types

- Tâches élémentaires et simples
- Jamais de tâches en parallèle
- Notion de gain de temps essentielle
- Exclue les tâches de recherche spatiale
- Exclue les processus demandant la visualisation d'informations

VUI versus GUI: Conception

■ Modes -> Choisir le meilleur

■ Auditif

- Info courte et simple
- Info ne doit pas être mémorisée
- Info temporelle
- Avertissement critique
- Réponse verbale et immédiate requise
- Visuel surexploité
- Environnement inapproprié pour un support visuel (mobilité, manque de lumière)

■ Visuel

- Info complexe et longue
- Info doit être mémorisée
- Info spatiale (ex: cartes)
- Auditif surexploité
- Environnement bruyant et/ou où faire du bruit est inacceptable

VUI versus GUI: Conception

■ Contraintes de conception

- Pas de formation ni d'apprentissage requis
- Nature dynamique:
Non persistance de l'information
- Multiples types d'utilisateurs
- Implication sociale de l'utilisateur
- Limites de la mémoire à court terme
 - 7 ± 2 items (Miller, 1956)
 - 3 ± 2 items dans les menus sans barge-in (Balantine, 2001)
- Personnalité de la machine



VUI versus GUI: Conception

■ Scénarisation

- Principale activité = écriture du dialogue
- Format en prose ou en dialogue
 - Conversations détaillées avec le système
- L'écriture doit se faire « oralement »
 - Pas les mêmes ressources mentales sollicitées en lecture qu'en écoute
 - Nature dynamique de l'interaction importante
- Focus sur les tâches principales, loi du 80/20
 - Décrire ce que 80% des gens feront 80% du temps.

VUI versus GUI: Conception

■ Scénarisation

■ Exemple de scénario:

Voice Activated Dialer :: Name Dialing

Scenario 0: Input OK + Phone type specified OK.

S: Operator!

U: Call <John Smith> at his <office>.	Connect (me?) to Call Give me (to at)? (His Her Their) Please	Work Office Mobile Cell Cellular Home Personal (line?) Pager
---------------------------------------	---	---

S: <John Smith> at <work>. Connecting you !

VUI versus GUI: Conception

■ Scénarisation

■ Exemple de scénario:

Voice Activated Dialer :: Name Dialing

Scenario 1: No Input + No phone type specified .

S: Operator!

U: - No input -

S: Say the name of one of your contacts, or give me his phone number and I'll dial it for you...

U: Call uh, John Smith!

S: For <John Smith> I have: mobile, home or work. Which one?

U: Uh, work !

S: <John Smith> at <work>. Connecting you !

VUI versus GUI: Conception

■ Guides de style -> GUI versus VUI

■ Fontes -> Mots

■ Couleurs -> Sons

■ Fonds d'écrans -> Ambiances sonores

■ Logos -> Signature sonore

■ Objets graphiques -> Phrases



VUI versus GUI: Conception

■ Challenges

- Le mode « conversation » (langage naturel) n'est pas nécessairement la **métaphore** idéale
- L'enjeu n'est pas seulement le prompt, mais aussi la **prosodie**:
 - On peut gagner en efficacité dans le choix de la terminologie mais aussi dans la prosodie
 - Ex: Plusieurs façons de lire « Voulez-vous Valider ou Annuler ? »
- Prévisibilité de l'interface
 - Nature temporelle des VUI ne permet pas de créer un modèle conceptuel fort



VUI versus GUI: Conception

■ Challenges

- Le recouvrement d'erreur est délicat et spécifique:
 - En cas de no input.
 - En cas de no match.
 - En cas d'environnement bruyant.
 - En cas de problèmes de connexion.

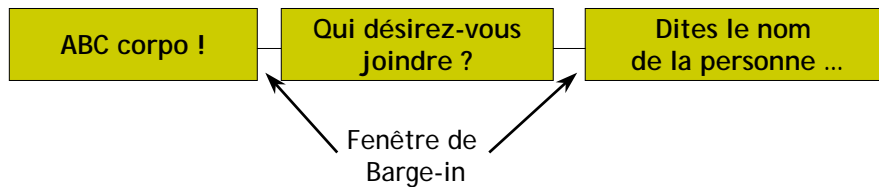


VUI versus GUI: Conception

■ Challenges

■ Le "turn-taking"

- Améliorer la prise de tour en cas de Barge-In.



Peut être une concaténation de segments coupés très courts, et espacés de blancs ou des pauses à l'intérieur d'un gros segment.



VUI versus GUI: Conception

■ Challenges

■ Output

- concaténation



- Continuous feed

- ATT Naturally Speaking:
- Eloquent:
- Fonix:

- Personnalité

- Earcons



VUI versus GUI: Évaluation

- Quelques distinctions ...
 - Heuristiques d'évaluation
 - Recueil de données
 - Comportements verbaux spontanés
 - Impossibilité du Penser tout haut
 - Tests d'utilisabilité
 - Setup physique du recueil de données
 - Classes de problèmes ergonomiques

VUI versus GUI: Évaluation

- Heuristiques d'évaluation - entre autres:
 - Contrôle
 - Visibilité de l'état du système
 - Clarté linguistique
 - Cohérence et standards
 - Prévention des erreurs
 - Limites humaines
 - Prévisibilité
 - Flexibilité
 - Esthétique et simplicité
 - Recouvrement d'erreur
 - Aide et assistance

VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Contrôle

- Contrôle: Offrir des façons de revenir sur une mauvaise action

- Confirm/Cancel



- Liberté de navigation lorsque possible

- Navigation dans une liste: Next, Previous, Start Over, Repeat.

- Pause ou clé #



VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Visibilité de l'état du système

- Sur un GUI le feedback peut être aussi subtil que d'indiquer par une dépression qu'un bouton est sélectionné.

- En VUI, le feedback DOIT être explicite. On doit, par exemple, confirmer explicitement qu'on a bien compris l'instruction que nous a donnée l'utilisateur; impose une interruption du flow.

- Earcon de confirmation aussi puissant que message verbal:



VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Clarté linguistique

- Le système devrait s'adresser à l'utilisateur dans son langage et ne pas employer de mots orientés système:
 - Vocabulaire du domaine
 - Vocabulaire de la tâche
 - Régionalismes si international
- Utiliser un ton et une terminologie qui conviendra aux types d'utilisateurs visés
 - Formel ou familier
 - Régionalismes
 - Jargon technique
 - Néologismes

Ex: Ticket se dit « boleto » au Mexique, « tiquete » en Colombie et « pasaje » en Argentine.

VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Cohérence et standards

- Cohérence dans la terminologie employée
 - « Start over » = recommencer le contenu d'une liste ou revenir au début de l'application ?
 - Si « Go Forward » est une commande acceptée, il faut alors que « Go Back » soit son corollaire et non seulement « Previous ».
- Cohérence inter et intra-application
 - Si « Main menu » = Revenir au menu principal du portail, ne pas le faire revenir au menu d'entrée d'un des services du portail ensuite.

VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Prévention des erreurs

- S'assurer que les utilisateurs savent exactement ce qui leur est demandé sans ambiguïté



VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Limites humaines

- Répéter les mots-clés à toutes les étapes
- Ne pas forcer à mémoriser, ne pas **présupposer** un apprentissage de la part des utilisateurs.
- Un output en TTS requiert plus de ressources et est plus difficile à décoder que la parole naturelle (Sanders & McCormick, 1993): donc ralentir l'output en TTS.

VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Prévisibilité

- Une interface doit s'apprendre par exploration, sans formation
- Elle doit permettre de prévoir comment l'application va réagir, donc de bâtir son modèle mental.



VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Flexibilité

- Relié aux types d'utilisateurs visés et aux tâches desservies
- Considérer le niveau de stress, l'interaction sociale
- Obéir au modèle de tâche
- Exemples:
 - Barge-in, Keyword barge-in, Hot keys
 - Archivage des données personnelles
 - Robustesse, rapidité d'exécution

VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Esthétique et simplicité

- Design attirant: Doit être la préférence des utilisateurs
 - Voix et ambiance sonore seront les premiers critères
- Design approprié: selon le domaine de la tâche, les utilisateurs peuvent préférer des voix synthétiques plutôt que naturelles
 - Ex: pilotes préfèrent les warnings en TTS (Cohen et Oviatt, 1994)
 - « *Making the simple complicated is commonplace; making the complicated simple, awesomely simple, now that's creativity.* » - Charlie Mingus

VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Recouvrement d'erreur

- Réduire le nombre et le coût des erreurs
- En cas d'incertitude de la reconnaissance:
 - Ne pas bloquer l'utilisateur, faire évoluer le dialogue
 - Offrir des options plus générales



Mauvais

Bon

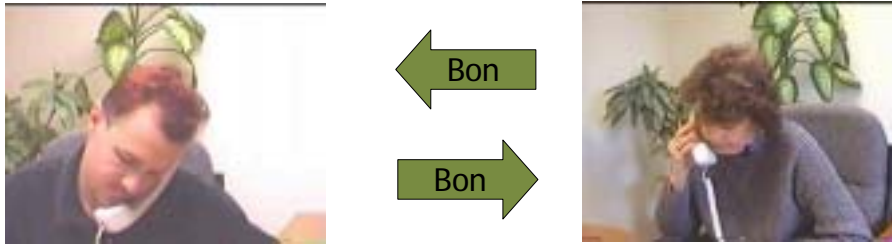


VUI versus GUI: Évaluation

■ Heuristiques d'évaluation

■ Aide et assistance

- Évolution du style de dialogue selon le type d'utilisateur
 - Diminution de l'assistance au fur et à mesure de l'interaction
 - De dialogue dirigé vers dialogue naturel



VUI versus GUI: Évaluation

■ Tests d'utilisabilité

- 10 utilisateurs représentatifs par session
- Papier-crayon impossible: exige un prototype robuste et fonctionnel
- Scénarios de tâches typiques et extrêmes
- Observer les comportements des utilisateurs puis procéder à des ajustements du call flow et du dialogue.
 - Erreurs, Temps d'exécution, Recours à l'aide, Satisfaction, etc.

VUI versus GUI: Évaluation

■ Recueil de données lors des tests

- Comportements verbaux spontanés
 - Charge mentale assez élevée
- Penser tout haut -> impossible
 - donc on n'a pas le raisonnement de l'utilisateur
- Interruption de l'interaction pour faire reprendre -> impossible
 - donc on doit déduire le cheminement pour arriver à une erreur
- **DONC**: il faut avoir un œil très exercé pour aller diagnostiquer les véritables problèmes.

VUI versus GUI: Évaluation

■ Conditions du recueil de données en tests

- En studio ou en mobilité ...
 - Caméra sur l'utilisateur
 - Pas d'image sur le dispositif
 - Prise de son ambiant
 - « Tapping » full duplex
 - Mixage des deux entrées en synchro sur la caméra



VUI versus GUI: Évaluation

■ Problèmes ergonomiques

■ Non détection d'erreurs d'état

- Le système fait une transition vers un autre état alors que l'utilisateur anticipe autre chose
- Le système ne fait pas de transition, mais l'utilisateur, lui a mentalement avancé.
 - "Je veux consulter mon portefeuille !"
 - "Voulez-vous effectuer une transaction ou consulter le solde d'un compte ?"
 - "Portefeuille !!!"

■ Faire avancer le dialogue à un stade où l'utilisateur a l'impression d'avancer.

VUI versus GUI: Évaluation

■ Problèmes ergonomiques

■ Non-détection des erreurs d'état: dialogue erroné

- "Mon portefeuille !"
- "Voulez-vous effectuer une transaction ou consulter le solde d'un compte ?"
- "Je veux mettre de l'argent dans mon portefeuille !"
- "Voulez-vous effectuer une transaction ou consulter le solde d'un compte ?"
- "Portefeuille !!!" ...

VUI versus GUI: Évaluation

■ Problèmes ergonomiques

■ Non-détection des erreurs d'état: dialogue révisé:

- "Mon portefeuille !"
- "Voulez-vous effectuer une transaction ou consulter le solde d'un compte ?"
- "Je veux mettre de l'argent dans mon portefeuille !"
 - *Ici: OOV, donc changement de grammaire et liste de choix plus vaste.*
- "Vous pouvez: Payer une facture, Consulter le solde d'un compte, ou transférer des fonds."
- "OK, transférer des fonds".

VUI versus GUI: Évaluation

■ Problèmes ergonomiques



Erreur d'état



Correction
d'erreur d'état

VUI versus GUI: Évaluation

- Problèmes ergonomiques
 - Listes et menus trop longs



VUI versus GUI: Évaluation

- Problèmes ergonomiques
 - Instructions trop longues
 - Instructions incomplètes



VUI versus GUI: Évaluation

- Problèmes ergonomiques
 - Feedback inadéquat ou absent
 - Terminologie inadéquate
 - Etc.



Interfaces vocales et Utilisabilité

- Méthodes de conception adéquates à cette technologie bien qu'ayant dû évoluer
 - Orientation autour du dialogue verbal et non visuel.
- Les professionnels ont dû aussi acquérir des connaissances supplémentaires sur le langage parlé et ses composantes et sur les technologies vocales.
 - Cela n'est pas, a contrario du langage visuel, enseigné à l'université !