



**HAL**  
open science

## Étiquettes vocales et utilisateur en situation d'illettrisme

Denis Chêne, Éric Petit, Tomas Pérez Segovia

► **To cite this version:**

Denis Chêne, Éric Petit, Tomas Pérez Segovia. Étiquettes vocales et utilisateur en situation d'illettrisme. Actes de la 31e conférence francophone sur l'Interaction Homme-Machine (IHM 2019), Dec 2019, Grenoble, France. pp.8:1-8, 10.1145/3366550.3372254 . hal-02381592

**HAL Id: hal-02381592**

**<https://hal.science/hal-02381592>**

Submitted on 26 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Étiquettes vocales et utilisateur en situation d'illettrisme

## *Vocal labels and user illiteracy*

**Chêne Denis**  
Orange labs.  
Grenoble, France  
denis.chene@orange.com

**Petit Éric**  
Orange labs.  
Grenoble, France  
eric.petit@orange.com

**Pérez Segovia Tomás**  
Orange labs.  
Grenoble, France  
tomas.perezsegovia@orange.com

### ABSTRACT

*Textual alternatives are insufficient to meet the needs of people with illiteracy and illectronism. We propose a new interface component named “Voice Labels” to overcome the difficulties of usability of people in such situations. User tests carried out on the ground in Burkina Faso are described. They highlight that the understanding of the interface impacts its manipulation, both at the level of the task (understanding of what it is necessary to do), and at the level of fine motor action (execution of tapping on the touch interface and finger sliding).*

### CCS CONCEPTS

• **Human-centered computing** → **HCI theory, concepts and models; Pointing; Visualization techniques.**

### KEYWORDS

*Universal design, Design for All, Vocal labels, Touch based interaction, Novice user, Illiteracy*

### RÉSUMÉ

Les alternatives textuelles sont insuffisantes pour répondre aux besoins des personnes en situation d'illettrisme et d'illectronisme. Nous proposons un nouveau composant d'interface nommé « Étiquette vocale » permettant de palier les difficultés d'utilisabilité des personnes dans de telles situations. Suite à des tests utilisateurs effectués sur le terrain au Burkina Faso, nous mettons en évidence que la compréhension de l'interface impacte sa manipulation, à la fois au niveau de la tâche

(compréhension de ce qu'il faut faire), mais également au niveau de l'action motrice fine (exécution des tapotements sur l'interface tactile et glissements de doigts).

### MOTS CLÉS

Conception universelle, Étiquettes vocales, Interaction tactile, Novice, Illettrisme

## 1 INTRODUCTION

Dans le domaine des interfaces accessibles aux personnes en situations de handicap la notion d'alternative textuelle est bien connue [20]. Elle permet entre autres de faciliter la vocalisation via des aides techniques telles que les « lecteurs d'écrans », ces derniers étant plus spécifiquement destinés aux personnes non-voyantes. Mais la vocalisation des interfaces n'est pas destinée uniquement à ce type d'utilisateurs. Les besoins des personnes illettrées sont du même ordre et vont même au-delà de cette problématique de vocalisation de l'interface. Les travaux de Medhi et al. [11][12] montrent que pour ce type de cibles, les besoins portent à la fois sur les types de représentations graphiques, la vocalisation des éléments de l'interface, l'aide fournie, la structuration du contenu, l'appropriation culturelle dont la langue n'est qu'une des composantes. Par exemple, des éléments d'interface représentés via un dessin plutôt que via une photo ou une représentation stylisée seront plus à même d'être compris. La photo est trop détaillée et pointe vers trop de détails inopportuns, la représentation stylisée est trop abstraite et par conséquent trop éloignée du réel et hors du cadre conceptuel des utilisateurs cibles. Le dessin par contre mixte les avantages d'une représentation à la fois concrète et limitant le niveau de détails tout en étant ancrée culturellement. Ces travaux soulignent également qu'une représentation audio, complémentaire de la représentation visuelle, est très efficace. Il s'agit donc à la fois de proposer des contenus adaptés mais également de les proposer sur les médias appropriés.

Dans la logique de conception universelle [4][6][17] la multimodalité de l'interface est un prérequis. De fait, la notion d'alternative est déclinée sous d'autres formes que le texte. En effet, une alternative peut être visuelle et animée dès lors qu'il s'agit de restituer la fonction par une vidéo en langue des signes (pour les utilisateurs pratiquant la langue des signes), ou bien graphique statique dès lors qu'il s'agit d'expliquer par un dessin la fonction proposée (pour les utilisateurs en situation de contrainte cognitive par exemple). Tout texte peut également être vocalisé (pour donner accès aux utilisateurs non-voyants ou illettrés). Pourtant, dans certaines langues, aucune synthèse vocale n'existe, et d'autres langues purement orales n'ont quant à elles même pas d'équivalence écrite. Ainsi les standards technologiques excluent la majorité des langues africaines [19]. Au final, les utilisateurs font face à des interfaces en langues étrangères et sont placés en situations d'illettrisme.

Le programme PIAAC [18] permettant de définir les compétences des adultes en termes de capacités de lecture dans des environnements numériques définit 6 niveaux d'illettrisme dont trois niveaux de base. Le premier est une incapacité totale à déchiffrer les textes, voire même à recopier une suite de chiffres. Le second correspond à la capacité de repérer quelques mots connus, et à se contenter d'un repérage des premières lettres telles que les majuscules alors que le troisième correspond à un niveau de déchiffrement de cours préparatoire. Mais l'illettrisme n'est pas qu'une question d'incapacité à lire du texte, il est également corrélé à des compétences cognitives dont l'implication sur la conception d'interface doit être prise en compte. De nombreuses études dans le domaine HCI4D<sup>1</sup> considèrent qu'il faut utiliser des interfaces combinant voix, vidéo et dessins, ainsi que d'autres compétences cognitives [1] liées à la maîtrise du langage, à l'efficacité [2], ainsi qu'à la gestion visuo-spatiale et aux capacités organisationnelles [3]. Katre [7] montre qu'il faut considérer que ces utilisateurs ont des structures cognitives et des mémoires linguistiques moins entraînées. Sherwani et al. [16] soulignent les difficultés de catégorisation de concepts abstraits et de navigation hiérarchique au sein d'un contenu. Cependant Medhi et al. [13][14] ainsi que Ladeira et al. [10] montrent que ces utilisateurs sont tout à fait aptes à apprendre des compétences présentées via des vidéos, et qu'une présentation des informations via des hiérarchies peu

profondes découpées en plusieurs écrans successifs peuvent mener à un usage acceptable. Un autre élément d'importance issu de ces études est que les utilisateurs peu lettrés ne sont pas seulement des consommateurs de contenus, mais peuvent également être des producteurs de contenus. Ceci a toute son importance dans le sens où de nombreuses interfaces non traduites nécessitent de l'être.

Une conception respectant toutes ces nuances socio-culturelles et socio-psychologiques devrait permettre à ces utilisateurs d'interagir de façon adéquate, si tant est que le dispositif d'interaction utilisé soit suffisamment manipulable. En effet les interfaces tactiles sont difficiles à utiliser pour les utilisateurs contraints [8][9]. Pour qu'une interaction puisse s'exécuter sereinement dans le cadre d'une perception ou d'une compréhension contraintes, il est nécessaire que les manipulations de l'interface puissent se faire de façon sécurisée. Ceci a été notamment été mis en œuvre et expérimenté Chêne et Petit dans le cadre de l'interaction tactile [5] [15].

Les interfaces à destination des personnes en situation d'illettrisme doivent résoudre l'ensemble de ces problématiques : des éléments d'interface simplifiés, organisés et présentés de façons claires et appropriées dans les contextes culturels adéquats, des manipulations sécurisées, ainsi que la possibilité de vocaliser les éléments d'interface y compris pour des langues uniquement orales. Nous testons ci-après une telle interface, enrichie d'un nouveau composant permettant d'associer aux fonctions de l'interface des étiquettes vocales auto-créées. Nous décrivons tout d'abord de dispositif expérimental, puis la mise en œuvre du test effectué, et discutons les résultats obtenus.

## 2 DISPOSITIF EXPÉRIMENTAL

### Prototype d'interface multimodale à étiquettes vocales

Les fonctions proposées par le prototype sont des fonctions de base de la téléphonie (lancer/recevoir un appel) mise en œuvre au sein d'un téléphone tactile. Il propose des rendus multimédias systématiques (vidéos, textes, pictogrammes, vocalisations, vibrations). Ces rendus sont déclinés sur des objets d'interface de même type (des listes de boutons), organisés dans une structure hiérarchique à faible profondeur. Les manipulations tactiles sont sécurisées et permettent à la fois d'empêcher les déclenchements inopportuns (erreurs de validations) issus d'effleurements d'interface non voulus, et de pointages approximatifs (glissements, écrasements, gros

<sup>1</sup>Human Computer Interface for Development

doigts, tremblements), mais aussi de manipuler l'interface tactile à un rythme lent via des validations en appui long et des manipulations continues (lorsque le focus est déplacé le doigt peut rester en contact avec l'écran sans limite de temps). Les phases de sélection et de validation sont dissociées. Un simple Tap génère une vocalisation de l'item en cours de focus. Il ne s'agit que d'une prise d'information sur l'interface, non d'une navigation. Un appui long statique de 2 secondes génère quant à lui la validation de l'item et la navigation subséquente au sein de la hiérarchie. Ces caractéristiques de manipulation lente et sécurisée permettent d'ouvrir un espace temporel propice à la prise d'information. Contrairement aux interfaces tactiles standards un tapotement de l'écran ne génère pas d'action de validation mais déclenche une « prise d'information » sur l'élément en cours de sélection. Dans le contexte d'une interface optimisée pour les personnes en situation d'illettrisme cette prise d'information est auditive. Chaque fonction peut ainsi être vocalisée. Le prototype de départ contient les libellés de fonctions en français et les « prises d'informations » déclenchent la synthèse vocale en français.

En sus, l'interface intègre une solution d'étiquetage à la volée des éléments d'interface. Ces étiquettes consistent en l'enregistrement d'une voix humaine associée directement aux items de l'interface. Un utilisateur bilingue (l'aidant), se positionne sur la fonction à étiqueter, lance l'enregistrement de sa voix, énonce la fonction après le bip d'enregistrement. Ainsi l'étiquette vocale est enregistrée et associée à la fonction. Cette étiquette prend le pas sur la synthèse vocale. De fait l'utilisateur illettré a désormais accès à une interface vocalisée y compris pour les langues ne disposant pas de système d'écriture. L'interrogation auditive de l'interface (sa vocalisation) se fait à l'aide d'un simple « tap » : dans un premier temps l'utilisateur sélectionne l'item qu'il souhaite interroger en plaçant le focus de sélection sur ce dernier, et dans un second temps il effectue un « tap » n'importe où sur l'écran. L'étiquette vocale correspondante est alors vocalisée par le système.

### Tests utilisateurs

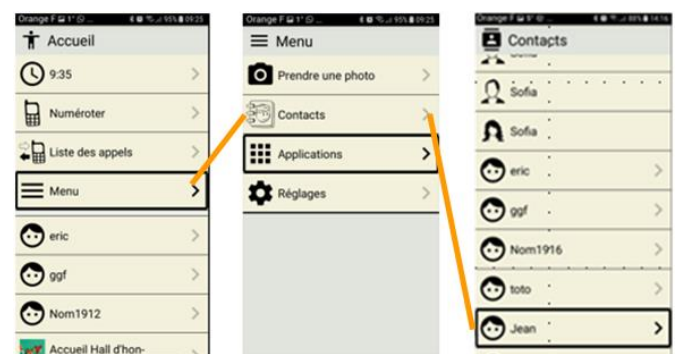
Des tests utilisateurs d'une telle interface tactile personnalisable ont été effectués par 10 personnes, 6 hommes et 4 femmes. 8 étaient en situation d'illettrisme de niveau 1, et 2 en niveau 2. En sus du questionnaire de recrutement, un exercice de recopie de chiffres sur l'écran tactile, nous a permis de définir l'appartenance à ces niveaux et de qualifier leur niveau initial de capacités de manipulation d'une interface tactile. Toutes étaient en

situation d'illectronisme et ne parlaient que le Mòoré (une des langues parlées au Burkina Faso). Elles possédaient toutes un téléphone non tactile mais ne l'utilisaient principalement que pour répondre aux appels. Toutes les passations se sont effectuées en Mòoré. Elles ont été menées avec interprète.

Leur tâche était d'accomplir les scénarios suivants (Figure 1), du plus simple au plus complexe : 1) identifier le contact « Toto » situé en page d'accueil, 2) trouver « Radio Ouaga » au niveau N-3 en début de liste des radios, 3) téléphoner au contact « Jean » situé au milieu de la longue liste des « Contacts » en niveau N-2 et enfin 4) identifier et lancer le service « Orange Money » situé au niveau d'arborescence N-5. Tout au long de l'expérimentation l'utilisateur pouvait traduire en Mòoré n'importe quel élément de l'interface.



Niveau N  
Scénario 1)



Niveau N

Niveau N-1

Niveau N-2  
Scénario 3)

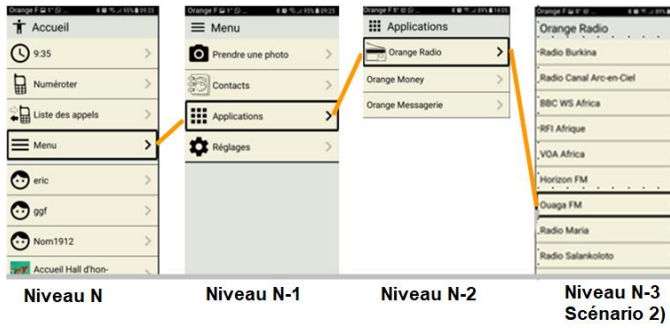


Figure 1: arborescence des scénarios 1) 2) et 3)

Nous avons recueilli le niveau de maîtrise du geste de tapotement sur une interface tactile. En fonction de leur qualité de production, nous avons distingué les **Manipulations Aisées** (tapotement bref, maîtrisé, déplacement posé du doigt, précision, utilisation intentionnelle de la délocalisation, verbalisations positives liées à la manipulation) des **Manipulations Contraintes** (tremblement, lenteur excessive, glissement, imprécision, hésitations, interrogations par des Tap non voulues, validations non voulues, inversion droite-gauche, verbalisations négatives liées à la manipulation).

Nous avons également relevé les réussites, semi-réussites ou échecs aux scénarios (Réussite / Semi-réussite / Échec). Une réussite consiste à atteindre l'objectif fixé sans aide de la part de l'expérimentateur. La semi-réussite est une atteinte de l'objectif avec aide. Un échec consiste à ne pas atteindre l'objectif fixé, malgré les aides apportées. Certains scénarios étaient exécutés très facilement, d'autres plus difficilement. Pour les qualifier nous avons codé les éléments indicateurs d'une **Tâches Aisées** (scénario résolu rapidement avec aisance, sans aide et associé à des verbalisations positives). Nous avons codé les éléments indicateurs de présence d'**Obstacles** (a minima des exécutions difficiles, et/ou des aides nécessaires, et/ou plusieurs essais-erreurs, et/ou de longues hésitations, et/ou des verbalisations négatives). Enfin nous avons codé pour chaque scénario le nombre de traductions effectuées. Lorsqu'aucune traduction n'était effectuée et que les utilisateurs ne s'interrogeaient pas sur la signification d'un terme, alors une note de **Traduction satisfaisante** était attribuée.

### 3 RÉSULTATS

#### Réussite / Échec aux scénarios

Les semi-réussites étant peu nombreuses (4 sur 40), nous avons cumulé les Réussites (36) et Semi-réussite (4) en une seule variable Réussites. En effet, les scénarios donnent

tous lieu à des réussites, pour tous les utilisateurs. C'est donc sur la qualité d'exécution des scénarios que nous avons travaillé. Le premier scénario 1) n'est pas exécuté très facilement et présente surtout des Obstacles à la tâche. Il est mêlé à la phase de découverte de l'interface et aucun signe n'indique qu'il est exécuté facilement. Il n'est pas codé « exécuté avec aisance » car une aide introductive est systématiquement apportée (notamment sur les commandes gestuelles à utiliser). Les utilisateurs sont tous hésitants en ce début de test. Le second scénario 2) présente quant à lui de nombreux indices d'une exécution facile. Les scénarios Jean et Orange Money sont plus difficiles à achever, certains utilisateurs cumulant même plusieurs Obstacles. D'autres cependant les exécutent avec facilité (Figure 2).

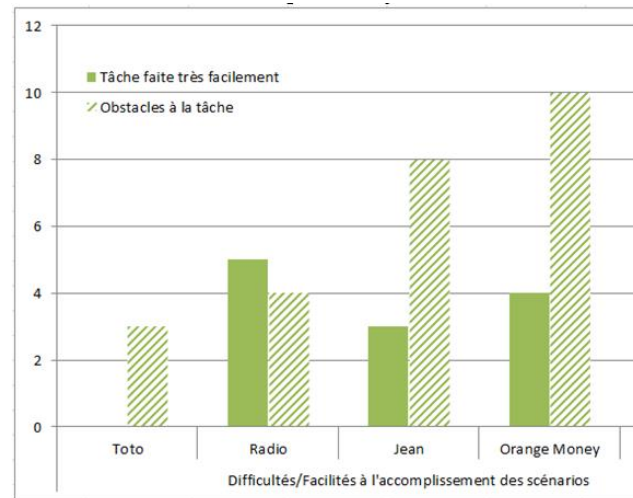


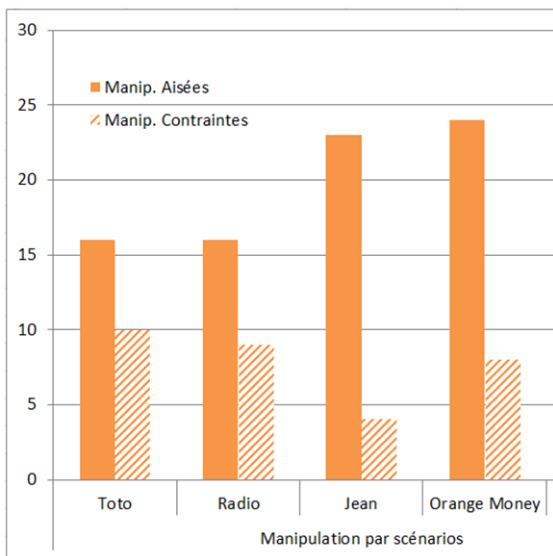
Figure 2: scénarios exécutés facilement vs scénarios présentant des obstacles

Les obstacles les plus communs sont « ne trouve pas le chemin », « ne trouve pas les contacts en fin de liste », « ne sait pas où aller », « ne comprend pas qu'il faut activer le bouton Continuer », « ne descend pas dans l'arborescence ». Malgré ces freins, suite à un encouragement ou de l'aide de l'expérimentateur, tous les utilisateurs ont réussi.

#### Manipulation fine de l'interface

Au-delà de la qualité des réussites aux scénarios, nous avons étudié la qualité des Manipulations tactiles de l'interface. Le geste de tapotement et plus généralement tout ce qui est de l'ordre de la manipulation fine de l'interaction sur écran tactile est de mieux en mieux maîtrisé par les utilisateurs au fur et à mesure que la passation avance (Figure 3). Les Manipulations Aisées sont significativement plus nombreuses que les Manipulations Contraintes. Une MANOVA confirme ce résultat au niveau

inférentiel  $F(4,15) = 7.4$ ,  $p < .005$ ; Wilk's  $\Lambda = 0.34$ . Il est à noter que dès le premier scénario les utilisateurs font preuve de Manipulations Aisées. Pourtant rappelons-le, ces utilisateurs sont novices en interaction tactile. Les Manipulations Aisées ou Contraintes ne sont pas différenciées au niveau du scénario Toto, puis elles le sont progressivement sur les scénarios « Radio » puis « Jean » et « Orange Money ».



**Figure 3: Le nombre de « Manipulations Aisées » vs les « Manipulations Contraintes » (occurrences de manipulations sur l'axe des ordonnées)**

### Traduction des interfaces

L'interface pouvait être traduite en Mòoré à tout moment, sur demande de l'utilisateur ou suite à identification d'une phase d'hésitation. Quelques termes ont été traduits au niveau de l'Accueil, une grande majorité de termes dans les listes du Menu et des Applications ainsi qu'au sein du menu d'Orange Money. Certains termes en Français n'ont pas nécessité de traduction (Message, Appeler...).

Les traductions ont ainsi évolué du début à la fin de chaque passation en fonction des incompréhensions mais aussi des caractéristiques des chacun. Certains ont par exemple proposé de remplacer le terme « Boutique » (en Mòoré) qu'ils comprenaient parfaitement, en « Marchand de Légumes » (en Mòoré) du fait de leur propre métier. Chaque utilisateur a ainsi pu personnaliser les intitulés de fonctions.

Les traductions ont été produites essentiellement par les quatre premiers utilisateurs. Les suivants ayant repris au maximum 2 à 3 traductions d'items qui leurs posaient problèmes.

La rubrique « Menu » (en Français) par laquelle il fallait passer pour atteindre le niveau N-1 a subi plusieurs phases de traductions. Ce terme ne possédant pas d'équivalent dans la langue cible il fut d'abord traduit en « Marché » (en Mòoré), puis en « lieu du marché, là où on peut trouver tout ce que l'on cherche » (en Mòoré). La rubrique « Applications » (en Français) donnant accès au service cible « Orange Money » (en Français) fut traduite d'abord en « liste des applications » (en Mòoré) puis en « Boutique » (en Mòoré) puis en « l'étal du marché, la boutique où il y a les produits précis que je cherche ». Enfin le terme « Réglages » (en Français) fut d'abord traduit en « Bouton réglage » (en Mòoré) puis en « réglages, outils du téléphone » (en Mòoré) du fait de la confusion possible avec le bouton de réglage de la radio.

En début d'expérimentation (pour les premiers sujets) les traductions sont nombreuses, l'interface étant jugée plus satisfaisante et moins sujette à traduction sur la fin de l'expérimentation ( $F(2,5) = 41.5$ ,  $p < .001$ ; Wilk's  $\Lambda = 0.06$ ). Ci-après nous explorons la relation entre ces traductions et la qualité de l'accomplissement des scénarios ainsi que sur la qualité des manipulations.

### Évolution de la Facilité d'accomplissement des scénarios vis à vis de la Traduction

Les premiers sujets ont éprouvé plus de difficulté que les derniers lors de l'exécution des scénarios. De plus l'interface nécessite davantage de traductions pour ces sujets. Les mesures de Facilité de la tâche et de Satisfaction des traductions diffèrent significativement entre le début et la fin de l'expérimentation ( $F(6,22) = 10.3$ ,  $p < .0001$ ; Wilk's  $\Lambda = 0.07$ ).

Les traductions sont davantage produites en début d'expérimentation, et les tâches sont exécutées plus facilement en fin d'expérimentation. En début d'expérimentation, la plupart des items n'étant pas en Mòoré, les obstacles à la tâche sont nombreux et les utilisateurs n'ont pas de comportements ou de verbatims exprimant que les scénarios sont faciles à accomplir. Mais en fin d'expérimentation, les traductions sont satisfaisantes et cela influe positivement sur la facilité d'accomplissement des scénarios (Figure 4).

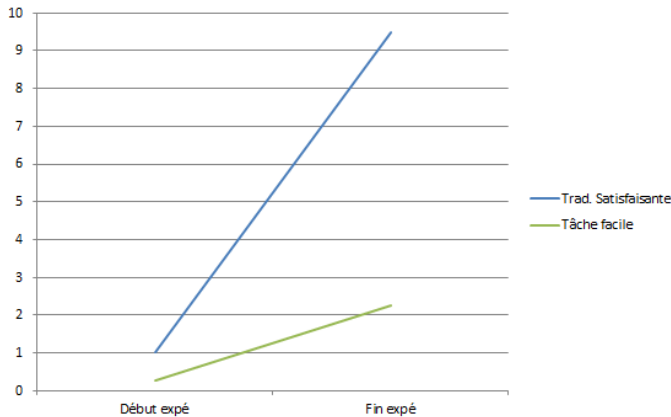


Figure 4 : interaction entre la mesure de Traductions satisfaisantes et Facilité d'accomplissement (moyennes)

### Évolution des Manipulations vis à vis de la Traduction

Plus l'expérimentation avance plus la traduction est satisfaisante, et plus les manipulations sont aisées (Figure 5). Nous observons une interaction entre le niveau de traduction et la qualité de la manipulation ( $F(6,22) = 8.4$ ,  $p < .0001$ ; Wilk's  $\Lambda = 0.09$ ). Au début de l'expérimentation les premiers utilisateurs ne sont pas très à l'aise en termes de manipulations et ne sont pas dans un environnement linguistique adéquat car il reste de nombreux termes à traduire. Par contre, en fin d'expérimentation ils sont dans un environnement connu et plus compréhensible et manipulent l'interface avec davantage d'aisance que les premiers utilisateurs.

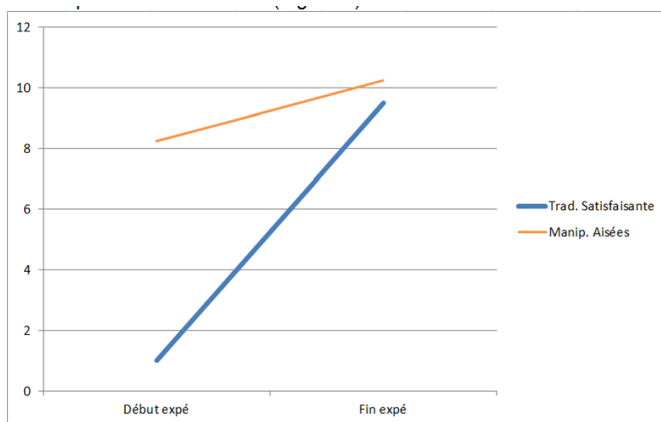


Figure 5 : interaction Manipulation/Traduction (moyennes)

## 4 DISCUSSION ET CONCLUSION

Les résultats de cette étude soulignent l'importance de la traduction d'une interface. Tout utilisateur se retrouve potentiellement en situation d'illettrisme lorsqu'il est

confronté à une interface en langue étrangère. Les personnes illettrées cumulent en sus des contraintes cognitives, des inhibitions liées au manque de confiance en soi, et l'absence d'entraînement aux manipulations requises sur une interface tactile. En effet, il n'est pas si simple de faire un simple Tap, ou même un double Tap. N'importe quel type d'utilisateur peut bénéficier des traductions effectuées, mais nous relevons de notre expérimentation que la qualité des traductions est en interaction avec la Facilité d'Accomplissement de la tâche ainsi que la qualité des Manipulations. Les gestes sont mieux ajustés à partir du moment où l'utilisateur anticipe leur trajectoire et leur cible dans un contexte personnalisé. La consultation des étiquettes vocales permet à ces utilisateurs de se conforter dans l'identification des items qu'ils consultent. En effet, dans les interfaces tactiles habituelles ils sont obligés de se lancer au hasard dans l'activation de l'item sans avoir eu d'information audio sur la nature de ce dernier. Bien sûr rien n'indique qu'il s'agit d'une relation de cause à effet, pourtant nous pouvons souligner que le fait de personnaliser les vocalisations de fonctions via des étiquettes vocales est un facteur facilitant l'apprentissage. L'utilisateur est confronté à une situation nouvelle de manipulations tactiles, d'arborescence de concepts peu familiers mais exprimés dans un langage qu'il comprend, sous une forme audio adaptée à ses faibles capacités de lecture. Les hésitations de manipulations sont moins fréquentes sur les derniers scénarios « Jean » et « Orange Money » (Figure 3). Cela traduit une prise de confiance et la levée des barrières de l'inhibition, propres à tout débutant. Une situation expérimentale comparant des interfaces vocalisées après exécution de l'action versus avant (comme dans le cas des étiquettes vocales) permettrait sans nul doute de confirmer ce point.

Le comportement des utilisateurs lors de ces tests confirme la tendance à se cantonner dans une interaction « à plat ». Lors des scénarios 2) et 3) ils explorent la liste de l'Accueil longuement de haut en bas avant de se décider à descendre dans l'arborescence des niveaux N-1, puis N-2 et N-3. Descendre dans l'arborescence est en effet une action à risque. L'utilisateur prend le risque de perdre ses repères. Cela nécessite la construction de catégories mentales et implique de solliciter la mémoire. Les utilisateurs les plus lents ont passé énormément de temps sur les listes de base (le niveau N de l'Accueil, et le niveau N-1 du Menu) avant de se décider à accéder aux sous-parties des items explorés. Ces observations confirment bien celles de Medhi et al. [14] sur la difficulté des personnes illettrées à naviguer en profondeur au sein

d'une interface. Néanmoins les réussites aux scénarios « Jean » et « Orange Money » (à hiérarchies nettement plus profondes) montrent que lorsque les manipulations sont sécurisées et qu'une interrogation audio personnalisée est disponible en amont de l'action, alors l'usage d'interfaces hiérarchiques devient possible.

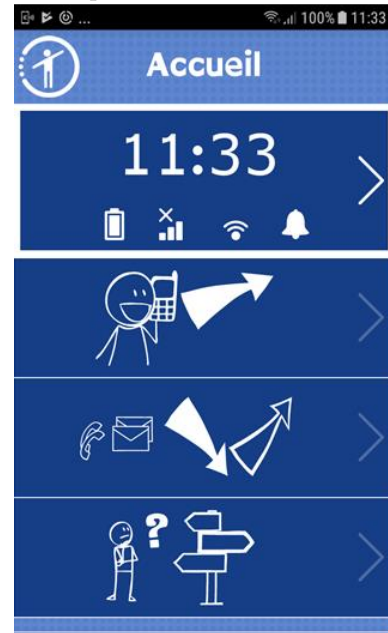
La manipulation sécurisée mise en œuvre, séparant la phase de sélection de la phase de validation ainsi que de larges zones tactiles pour les boutons et des délais d'activation suffisants, contribue grandement à la facilitation de la navigation au sein de l'interface. La prise en main de l'interface est bien meilleure sur les derniers scénarios « Jean » et « Orange Money », et les derniers utilisateurs, bénéficiant d'une traduction plus aboutie sur l'ensemble des scénarios, sont eux-mêmes plus à l'aise dans leurs manipulations.

Navigation sécurisée et exploration via les étiquettes vocales se combinent, permettant une compréhension progressive des items comme de la hiérarchie des fonctions.

Soulignons également que la nature même des traductions a évolué au cours des tests. Lorsque le terme à traduire avait son équivalent dans la langue cible, l'utilisateur traduisait d'un ou deux mots équivalents. Par contre, lorsque le concept n'existait pas c'est plutôt un groupe de mots, voire même une phrase complète qui faisait office de traduction. C'est un des avantages des étiquettes vocales : elles ne sont pas freinées par les limites d'affichage. En effet généralement seul l'affichage d'un ou deux mots peuvent définir une fonction du fait du manque de place sur l'écran. Mais lorsqu'il s'agit de faire comprendre à l'utilisateur un concept numérique ne faisant pas partie de sa langue, le libellé de la fonction n'est pas suffisant. C'est pourquoi de nombreuses étiquettes vocales se sont ainsi transformées en phrases d'explications. Ce résultat permet de réinterroger la subtile nuance entre le traditionnel libellé visuel de la fonction, a priori court puisque devant être affiché, et l'aide contextuelle de la fonction, plus longue et proposée dans un autre objet d'interface telle qu'une infobulle (audio ou visuelle). Dans notre expérimentation l'étiquette vocale est bien un libellé audio de fonction, mais c'est un libellé long au potentiel explicatif bien plus important.

À partir du moment où l'explication peut être enrichie sur le plan auditif la présentation du texte de la fonction ne semble pas nécessaire pour les utilisateurs illettrés de niveau 1. Il serait ainsi intéressant d'investiguer l'effet résultant de la disparition de l'affichage du texte décrivant les fonctions car après tout il est fort probable que pour ces utilisateurs le texte ne soit qu'un élément distracteur.

Nous projetons ainsi de tester et comparer à celle de cet article une interface 100% pictographique enrichie via des étiquettes vocales (Figure 6). Cependant une telle mise en œuvre nécessite avant tout de travailler sur la contextualisation/localisation des pictogrammes utilisés. En effet, nous l'avons vu, les termes utilisés pour les fonctions (tel que le terme Menu) ne sont pas toujours présents dans les langues cibles. Or les pictogrammes utilisés en Europe sont standardisés et calqués sur des normes culturelles numériques a priori très différentes de celles utilisés dans certains pays d'Afrique. Il convient avant toute chose de construire un catalogue de pictogrammes adéquats dans ce contexte culturel.



**Figure 6 : Niveau N de l'interface en mode 100% pictographique (Heure/Communiquer/Journal des appels/Menu)**

Ces travaux d'interaction homme-machine se poursuivent dans le cadre des principes d'une conception universelle axée sur les interfaces multiprofiles et adaptatives. Ainsi, au-delà des questions de traduction des interfaces, nous cherchons à adresser différents types de contraintes : celles liées à l'exécution des commandes tactiles, celles liées aux perceptions visuelles et/ou auditives des rendus de l'interface, et celles liées aux capacités de compréhension des contenus présentés. Ainsi nous projetons de créer et tester d'autres profils d'utilisation basés sur les contraintes motrices telles que la capacité de n'exécuter qu'un seul et unique geste de commande (mode Moteur++), la capacité à n'effectuer que deux ou trois gestes différents (mode Moteur+), ou encore à n'effectuer que des microgestes de pointage (mode



Microgestes). D'autres profils d'utilisation sont orientés sur les contraintes perceptives, avec par exemple la capacité de ne percevoir l'interface que de façon auditive (mode Vocal+), ou bien de la percevoir visuellement de façon imparfaite et complétée d'informations auditives et tactiles (mode Vision+). Chaque mode se traduit par une combinaison multimodale de commandes et rendus. Ces modes d'interaction personnalisés sont la clé de voûte de l'approche de conception universelle.

## REMERCIEMENTS

Nous remercions tout particulièrement les utilisateurs qui ont participé au test, ainsi que les membres du Ouagalab, haut lieu d'innovation à Ouagadougou.

## RÉFÉRENCES

- [1] Abadzi, H. 2003. Improving adult literacy outcomes: Lessons from cognitive research for developing countries. *Washington, DC: World Bank*.
- [2] Ardila, A., Rosselli, M., & Rosas, P. 1989. Neuropsychological assessment in non-literates: Visuospatial and memory abilities. *Brain Cognition*, 11(2), 147–166.
- [3] Bandura, A. 2005. Self-efficacy. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (pp. 71–81). New York: Academic Press
- [4] Chêne, D., Petit, É., & Zijp-Rouzier S. 2016. How to Achieve Design for All. Chapter · July 2016. DOI: 10.1007/978-3-319-40244-4\_12, In book: *Universal Access in Human-Computer Interaction. Interaction Techniques and Environments*, pp.117–128 (HCII 2016).
- [5] Chêne, D., Pillot, V., & Bobilier-Chaumont, ME. 2016. Tactile Interaction for Novice User – Uncollocated Gestures, In Springer Ed. *Volume 24, LNCS\_9754*.
- [6] Chêne, D., Botherel, V., Joucla, H. 2019. Une conception universelle mise en œuvre via des modes d'usages. Ed. Le Sensolier. *Actes de la 17<sup>ème</sup> journée du Sensolier, 2 oct. 2019, Paris*.
- [7] Katre, D. S. 2006. Unorganized cognitive structures of non-literate as the key factor in rural e-learning design. *I-Manager's Journal of Education Technology*, 2(4), 67–72.
- [8] Hwangbo, H., Yoon, S.H., Jin, B.B., Han, Y.S., Ji, Y.G.: A study of pointing performance of elderly users on smartphones. *Int. J. Hum. Comput. Interact.* 29(9), 604–618 (2013).
- [9] Kobayashi, M., Hiyama, A., Miura, T., Asakawa, C., Hirose, M., Ifukube, T.: Elderly user evaluation of mobile touchscreen interactions. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) *INTERACT 2011, Part I. LNCS*, vol. 6946, pp. 83–99. Springer, Heidelberg (2011).
- [10] Ladeira, I., and Cutrell, E. 2010. Teaching with storytelling: An investigation of narrative videos for skills training. *4th ACM/IEEE International Conference on Information and Communication Technologies and Development, IEEE, London, December, 2010*.
- [11] Medhi, I. & Cutrell, Ed. & Toyama, K. 2010. It's not just illiteracy. *India HCI in conjunction with the IFIP TC13 Special Interest Group on Interaction Design for International Development, 2010*.
- [12] Medhi, I., Patnaik, S., Brunskill, E., Nagasena Gautama, S.N., Thies, W. & Toyama, K. 2011. Designing Mobile Interfaces for Novice and Low-Literacy Users. *ACM Transactions on Computer-Human Interaction* 18(1):2 (TOCHI 2011).
- [13] Medhi, I., Lakshmanan, M., Toyama, K. and Cutrell, E. 2013. Some Evidence for Impact of Limited Education on Hierarchical User Interface Navigation. *ACM Conference on Human Factors in Computing Systems, 2013*.
- [14] Medhi, I., Toyama, K., Joshi, A., Athavankar, U. & Cutrell, E. 2013. A Comparison of List vs. Hierarchical UIs on Mobile Phones for Non-literate Users. *Chapter from book Human-Computer Interaction – INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part II pp.497-504*.
- [15] Petit, É., Chêne, D. 2015. MenuDfA : vers une navigation gestuelle tactile conçue pour tous. DOI : 10.13140/2.1.4929.2164, 2015.
- [16] Sherwani, J., Palijo, S., Mirza, S., Ahmed, T., Ali, N. and Rosenfeld. R. 2009. Speech vs. Touch-tone: Telephony Interfaces for Information Access by Low Literate Users. In *Proc. Information & Communications Technologies and Development, Doha, Qatar, April 2009*.
- [17] Stephanidis, C., Akoumianakis, D. & Savidis, A. 2001. Universal Design in Human-Computer Interaction. In book: *International Encyclopedia of Ergonomics and Human Factors, Vol. 1, Publisher: Taylor & Francis, Editors: W. Karwowski, pp.741-745, 2001*.
- [18] The Program for the International Assessment of Adult Competencies (PIAAC). OECD program. The Organisation for Economic Co-operation and Development (OECD). <http://www.oecd.org/>
- [19] Unesco, Mesurer la diversité linguistique sur l'Internet, Paris, Publications Unesco, 2005.
- [20] Web Accessibility Initiative. 2014. Resources on Alternative Text for Images. <https://www.w3.org/WAI/alt/>