
INTERACTION GESTUELLE AVEC LES SYSTÈMES EMBARQUÉS

Jean-François Kamp et Franck Poirier

Laboratoire VALORIA
Université de Bretagne-Sud
Rue Yves Mainguy 56000 Vannes
(Jean-Francois.Kamp, Franck.Poirier)@univ-ubs.fr

Philippe Doignon

Renault – Technocentre
1, avenue du golf
78288 Guyancourt Cedex
philippe.doignon@renault.fr

Résumé

Cet article présente une étude réalisée en collaboration avec la Direction de la Recherche de Renault. L'étude s'intéresse plus particulièrement à l'interface d'entrée des systèmes interactifs installés dans l'habitacle du véhicule. Un des problèmes majeurs de l'interface d'entrée est lié à la limitation des commandes actuelles qui se réduisent souvent à de simples boutons. L'interaction est d'autant plus difficile que d'une part, les nouveaux systèmes (navigation, téléphone) sont difficiles à utiliser et que d'autre part, le conducteur est pris par sa tâche principale de conduite qui implique une charge cognitive importante. Pour résoudre ce problème, une nouvelle interface tactile de commande est proposée et analysée. La surface tactile, combinée à un système de reconnaissance, permet de capter et de reconnaître n'importe quel tracé (lettres, chiffres, symboles) effectué par le doigt de l'utilisateur. L'analyse se décompose en deux parties. Dans la première, nous montrons que toute personne est capable de former correctement des lettres majuscules sans bénéficier d'aucun retour visuel. Dans ces conditions, le taux moyen de reconnaissance s'élève à 93 %. La seconde partie est consacrée à une définition d'un vocabulaire de symboles beaucoup plus riche que celui offert par de simples boutons. Une première étape encourageante est franchie pour prouver que le dispositif proposé rend l'interaction personnes-systèmes vraisemblablement compatible avec la tâche de conduite. Ce dispositif permet, en outre, à l'utilisateur de dialoguer efficacement avec les systèmes interactifs.

1. Introduction

Mis à part l'autoradio disponible en option depuis longtemps, l'interface Homme-machine en véhicule se limitait jusqu'il y a peu à de "simples" boutons et manettes en entrée et à quelques cadrans (compteurs, jauges) et voyants lumineux en sortie.

Depuis la fin des années 80, des systèmes électroniques plus sophistiqués ont été développés sous l'impulsion de grands projets européens ("PROMETHEUS", PROgram for a European Traffic system with Highest Efficiency and Unprecedented Safety, (Renault, 1995), "DRIVE", Dedicated Road Infrastructure for Vehicle safety in Europe, (Leiser, 1991a)), américains ("IVHS", Intelligent Vehicle Highway System, (Catling, 1994)) ou japonais ("RACS", Road/Automobile Communication System, "AMTICS", Advanced Mobile Traffic Information). Ces nouveaux systèmes ont tous pour objectif commun l'amélioration du trafic routier et de la sécurité. Mais là où les technologies récentes permettraient des progrès impressionnants (guidage automatique du véhicule, vision améliorée, système anti-collision), souvent les concepteurs se heurtent à un problème d'interface Homme-machine lors de la phase d'installation du nouveau système en véhicule. En effet, les règles suivies pour assurer une ergonomie optimale s'inspirent souvent de notions uniquement applicables aux interfaces entre l'Homme et l'ordinateur statique. Or, leur

transposition à un contexte mobile et plus particulièrement au domaine automobile est, dans bien des cas, tout à fait inadaptée.

Cet article est organisé en quatre parties. Au second paragraphe, les spécificités d'une interaction Homme-machine en véhicule sont présentées et la notion de "système embarqué" est précisée. Ensuite, la problématique de l'interaction avec les systèmes embarqués est examinée et une solution novatrice, par rapport au contexte automobile, est proposée : il s'agit de la modalité gestuelle. Au paragraphe quatre, une étude est menée en laboratoire afin d'analyser la compatibilité de la modalité gestuelle avec la conduite automobile. Finalement, un vocabulaire gestuel est défini pour permettre à l'utilisateur de dialoguer efficacement avec les systèmes embarqués.

2. Interaction Homme-machine en véhicule

2.1. La conduite automobile

Le conducteur d'un véhicule automobile effectue une tâche prioritaire sur toutes les autres qui est la conduite de son véhicule. A elle seule, cette tâche sollicite fortement l'individu au niveau sensori-moteur. Des cinq sens de l'être humain, quatre sont spécialement concernés :

- La vue : l'information visuelle est très dense et provient à la fois de l'extérieur du véhicule (trafic, signaux routiers) et de l'intérieur (voyants lumineux, afficheurs).
- L'ouïe : l'oreille capte les sons extérieurs à l'habitacle (coup de klaxon, freinage brusque d'un autre véhicule), et les bruits émis par le véhicule sont également porteurs d'informations (le changement de vitesse s'effectue au moment où le régime moteur est élevé).
- Le toucher ou plus généralement le canal tactilo-proprio-kinesthésique¹ : les vibrations et l'inclinaison du véhicule, la forme des boutons et des manettes, le cliquetis d'une molette, la dureté d'une pédale sont des informations perceptibles par l'être humain.
- l'odorat : l'odeur agit surtout sur le confort des occupants du véhicule mais il peut aussi transmettre de l'information : " le moteur sent le brûlé ".

La conduite automobile est une tâche complexe. Il est d'usage (Neboit, 1994) de la structurer en trois sous tâches qui se déroulent en parallèle :

- une tâche opérationnelle c'est-à-dire le contrôle de la trajectoire et la régulation de la vitesse ;
- une tâche tactique qui correspond au guidage du véhicule dans le réseau routier (choix de la voie, dépassement) ;
- une tâche de navigation ou de planification de l'itinéraire.

Pour accomplir cette tâche de conduite du véhicule, le conducteur dispose de commandes indispensables telles que le volant, les pédales ou la manette des clignotants.

2.2. Systèmes embarqués

Parallèlement à la tâche prioritaire de conduite, l'individu interagit également avec des systèmes électroniques :

- pour augmenter sa capacité de communication avec le monde extérieur (téléphone, consultation de données distantes) ;
- pour améliorer sa sécurité (système anti-collision, vision améliorée) ou son confort (autoradio, climatisation) ;
- pour faciliter ses déplacements (système de navigation ou d'information trafic).

Ces systèmes électroniques sont appelés " systèmes embarqués ". Un système embarqué possède les caractéristiques suivantes :

¹ Le toucher fait référence principalement aux mains et aux pieds. Le canal tactilo-proprio-kinesthésique englobe également toute l'enveloppe corporelle, les couches musculaires, les récepteurs articulaires et vestibulaires.

- il n'est pas indispensable à la conduite du véhicule ;
- il n'est pas spécifique au milieu automobile mais il est adapté pour une utilisation en véhicule. L'autoradio par exemple est une radio étudiée pour être logée dans l'habitacle ;
- il offre généralement un grand nombre de fonctionnalités. Le téléphone par exemple permet à la fois de composer un numéro et de gérer un répertoire téléphonique.

L'interaction avec les systèmes embarqués se fait par l'intermédiaire de commandes autres que celles nécessaires à la conduite. **Pour une question évidente de sécurité, l'utilisation de ces commandes ne doit pas augmenter la charge cognitive qui pèse déjà sur le conducteur.** Le code de la route n'interdit pas l'usage des systèmes embarqués mais il impose que " *Tout conducteur de véhicule doit se tenir constamment en état et en position d'exécuter commodément et sans délai toutes les manœuvres qui lui incombent (Article R.3-1).* "

L'étude exposée dans cet article concerne essentiellement l'interface d'entrée des systèmes embarqués, c'est-à-dire la commande. L'interface de sortie (affichage graphique, informations auditives) est peu abordée même s'il s'agit également d'une thématique de recherche (Ouadou *et al.*, 1993 ; Pleczon *et al.*, 1994) qui a son importance.

3. Problématique et solution proposée

Le premier paragraphe de cette section décrit quelques interfaces de commande traditionnelles des systèmes embarqués. Une courte analyse permet à chaque fois de faire ressortir les défauts de ces commandes.

Ensuite, l'interaction vocale est abordée succinctement. Les points forts de cette approche sont brièvement présentés.

Au troisième paragraphe, nous proposons une solution alternative à la modalité parole : il s'agit de la commande gestuelle qui est décrite en détail.

3.1. Problématique

Les dispositifs de commande les plus couramment installés dans l'habitacle sont les boutons et les croisillons (Figure 1). Dans la suite, un croisillon sera considéré comme un bouton. Ces dispositifs sont très mal adaptés à une utilisation en véhicule pour diverses raisons (Franzen, 1991) :

- ils sont trop nombreux et l'utilisateur associe difficilement un bouton à une action ;
- en conduisant, il est pratiquement impossible de les manipuler sans détourner le regard ; les boutons répondent donc très mal au besoin de sécurité ;

- les fonctions réalisées par les boutons sont assez pauvres : ils permettent uniquement d'ajuster une valeur (volume sonore par exemple), de sélectionner une rubrique d'un menu, d'annuler une action ou de parcourir une liste.
- bien souvent, l'accessibilité aux boutons est mauvaise parce qu'ils sont trop petits et situés trop loin des mains et du volant.

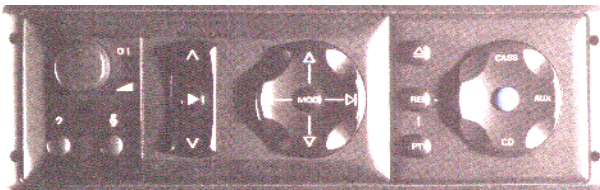


Figure 1 – Les boutons et les croisillons

D'autres systèmes de commande tentent d'améliorer l'accessibilité et le problème d'encombrement des boutons. Le " bouton rotatif " décrit à la Figure 2 en est un exemple. Celui-ci est monté sur la planche de bord. Il se présente sous la forme d'un tronc de cône crénelé que l'on actionne en tournant dans un sens ou dans l'autre, mais également en appuyant comme un bouton poussoir pour valider. Le mouvement de rotation du bouton permet à l'utilisateur de déplacer un curseur sur un écran graphique suivant un axe. Un autre bouton poussoir permet à l'utilisateur d'annuler sa dernière action.

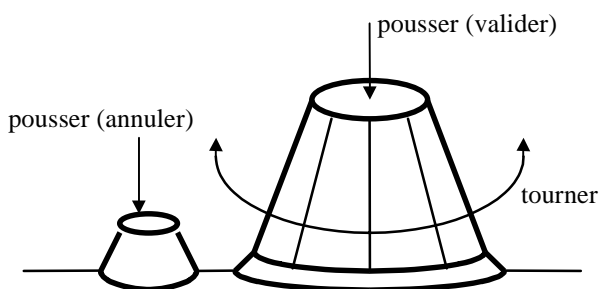


Figure 2 – Le schéma du bouton rotatif

Ce système de commande est facile d'accès s'il se situe à proximité du volant et regroupe un nombre limité de boutons. Il convient bien lorsque les manipulations à effectuer restent simples : sélection de quelques rubriques dans une arborescence de menus peu profonde. Par contre, il devient rapidement limité et peu performant lorsqu'il est nécessaire d'interagir avec des dispositifs embarqués plus complexes. En particulier :

- Lorsqu'il s'agit de désigner une zone sur une carte (système de navigation), la rotation n'autorise un

déplacement que suivant un seul axe alors que la désignation se fait dans un plan.

- Le choix d'une destination pour un système de navigation (nom de la ville, nom de la rue) ou l'inscription d'un nouveau correspondant dans le répertoire téléphonique, nécessitent l'entrée de lettres et de chiffres. Cette tâche peut être réalisée en tournant le bouton rotatif et en parcourant séquentiellement l'alphabet. En appuyant sur le même bouton la lettre désirée est sélectionnée. Il apparaît clairement que dans ce cas la manipulation du bouton rotatif est longue et fastidieuse puisque pour chaque lettre qui compose un mot il faut à nouveau parcourir l'alphabet.
- Bien que l'accès à la commande soit aisé (un simple repérage tactile suffit à identifier le bouton rotatif du bouton " annuler "), presque chaque manipulation de la commande nécessite une déportation du regard sur un écran graphique pour vérifier son résultat (parcours de l'alphabet par exemple).

3.2. Autre approche : la parole

La modalité parole semble être, a priori, une solution bien adaptée au problème d'interaction avec les systèmes embarqués, pour diverses raisons :

- l'homogénéité entre l'interface d'entrée (reconnaissance de parole) et l'interface de sortie (synthèse de parole) est parfaite puisque dans les deux cas il s'agit de la voix ;
- comme évoqué précédemment, d'un point de vue sécurité de conduite, la communication par la parole évite les détournements du regard et permet de garder les mains sur le volant (Leiser, 1991b) ;
- les dispositifs d'interaction sont simplifiés car limités à quelques micros en entrée et des haut-parleurs en sortie, ce qui allège considérablement l'interface matérielle ;
- la structure arborescente habituelle constituée de menus et sous-menus devient beaucoup plus horizontale car la reconnaissance de la parole couplée à un système de compréhension du langage naturel (Price, 1995) permet à l'utilisateur de formuler sa requête spontanément en une seule fois.

L'usage de la parole en véhicule présente néanmoins certains inconvénients :

- le vocal ne peut pas être utilisé pour accomplir n'importe quelle tâche (Néel et al., 1996) : déplacer un curseur sur une carte n'est pas aisé en mode vocal ;
- la reconnaissance et la synthèse de parole peuvent, malgré tout, augmenter la charge cognitive (obligation de respecter une grammaire, synthèse vocale de mauvaise qualité, ...)

- la synthèse de parole n'est pas forcément bien acceptée par l'utilisateur : un système "trop bavard" fatiguera rapidement le conducteur (Leiser, 1993), l'interaction doit être réduite à des messages brefs et synthétiques (Néel et al., 1996) ;
- la reconnaissance de parole embarquée est un système complexe et onéreux qui se compose globalement d'un réducteur de bruit ambiant — imposant l'installation de plusieurs micros (Smolders et al., 1994) —, d'un reconnaisseur de mots enchaînés et d'un analyseur de langage naturel.

3.3. Solution proposée : la modalité gestuelle

Le geste est une nouvelle modalité d'interaction avec les systèmes embarqués. Elle possède de nombreux avantages comme nous le verrons plus loin. Le but de l'étude n'est pas d'opposer les commandes orales et gestuelles. Il s'agit plutôt de consacrer l'étude à une analyse en profondeur du mode gestuel pour mettre en évidence ses qualités.

Le geste de l'utilisateur est capté par une surface tactile. La surface tactile utilisée au cours de notre étude se présente comme une tablette de faible épaisseur et d'environ 6 cm de côté. Sa résolution est de 118 dpi². Tant que le doigt exerce une pression suffisante sur la surface, un point de coordonnées (x, y) est transmis par liaison série à une unité de traitement. De cette manière, n'importe quel geste (ou tracé) du scripteur est capté par le dispositif pour être traité et reconnu. Etant donné qu'il n'est pas toujours possible de disposer d'un outil d'écriture dans le véhicule, le conducteur est amené à utiliser son doigt pour dessiner un symbole. Le mot "geste" étant un terme assez général et souvent dépendant du système de capture, il est défini comme étant, dans notre cas, le déplacement du doigt au contact de la surface tactile qui conduit à un tracé quelconque pouvant représenter une lettre, un chiffre, un symbole ou un pictogramme.

L'usage d'une surface tactile en véhicule a déjà fait l'objet de différentes analyses (Kalisvaart et Hartemann, 1991 ; Coleman et al., 1991). Dans tous les cas, la surface tactile, qui est transparente, est placée devant un écran de visualisation de même dimensions. Par configuration logicielle, l'écran fait apparaître des zones (un clavier AZERTY, des boutons, des icônes) en arrière plan du pavé tactile. L'utilisateur désigne alors ces différentes zones sensibles avec le doigt et la position du doigt est identifiée par la surface tactile. L'avantage principal est la possibilité de configurer les zones sensibles "à volonté" et suivant le contexte : par exemple, le clavier AZERTY n'apparaîtra que lorsque l'application nécessite l'entrée de données alphanumériques. Cette configuration permet une

interaction beaucoup plus riche que celle proposée par les boutons. Malgré tout, ce système de commande présente un certain nombre d'inconvénients :

- les analyses suggèrent des zones sensibles de grandes tailles (2,6 cm de côté (Sears, 1991)) pour limiter les erreurs de pointage. Ceci est contradictoire avec la taille de la surface tactile qui doit rester limitée en véhicule,
- si les zones sont petites, le pointage avec le doigt oblige un détournement du regard surtout lorsque l'écran est placé loin du volant. Or, il est important de limiter ces détournements surtout en conduisant,
- les actions de pointage pour l'accomplissement de tâches complexes demandent des déplacements répétés du bras et de la main entraînant rapidement une certaine fatigue (Bellik, 1995).

Tout en préservant l'avantage d'une interaction riche en terme de taille du vocabulaire, la nouveauté de notre étude consiste à déporter la surface tactile par rapport à l'écran d'affichage. De cette façon, la surface sensible peut aisément se placer à proximité des mains et du volant et il n'est plus question de l'utiliser comme dispositif de pointage. En effet, le but est de reconnaître le geste de commande tracé n'importe où avec le doigt sur la surface. Il n'y a donc plus de lien direct entre le pavé tactile et l'écran graphique. À partir du moment où un système reconnaît le geste tracé sans ambiguïté, il est possible d'y associer une action spécifique qui peut être de différents types :

- une action qui permet de plus facilement circuler dans l'arborescence des menus de l'application : revenir au menu précédent, retourner à la racine, passer au menu suivant,
- une action qui autorise l'utilisateur à accéder directement à une prestation sans suivre obligatoirement les chemins imposés par une arborescence : tracer la lettre "T" pour entrer dans la prestation téléphone, "N" pour le système de navigation. De manière équivalente à la notion de "raccourci clavier" pour l'ordinateur (Shneiderman, 1998), nous définissons ici la notion de "raccourci gestuel" pour les systèmes embarqués,
- la reconnaissance d'un caractère alphanumérique (lettre et/ou chiffre) dans le but :
 - soit d'accéder plus rapidement à un élément alphanumérique appartenant à une liste ou une base de données,
 - soit d'écrire un nouvel élément dans une liste.
- une action de base qui aide l'utilisateur dans l'accomplissement de sa tâche : annuler une commande, valider, demander de l'aide,

² " Dots Per Inch " en Anglais.

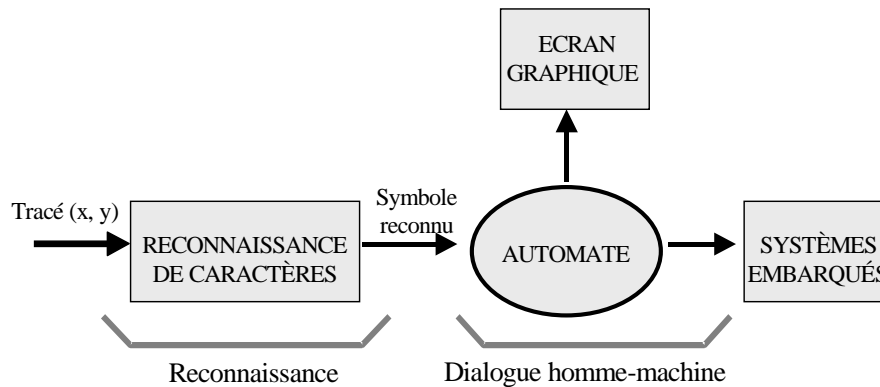


Figure 3 – Le système de commande tactile dans son ensemble

- une action spécifique dépendante du contexte comme par exemple : effectuer un zoom sur la carte, dérouler la liste des parkings disponibles.

Tout l'intérêt de la surface tactile couplée à un système de reconnaissance de gestes (caractères) réside donc dans :

- la liberté de définir n'importe quel symbole gestuel et d'y associer une action précise,
- la possibilité de remplacer par le tracé d'un seul symbole, les longues manipulations de la molette et des boutons pour effectuer la même action.

La Figure 3 représente le principe de commande tactile des systèmes embarqués. Celui-ci peut se décomposer en deux modules :

1. Un premier module appelé "Reconnaissance", destiné exclusivement à reconnaître le tracé effectué avec le doigt sur la surface tactile,
2. Un second module, appelé "Dialogue Homme-machine" spécifique au dialogue avec les systèmes embarqués. La Figure 3 montre que le dialogue est géré par un automate et que l'information est fournie à l'utilisateur par l'intermédiaire d'un écran graphique.

L'article expose aux paragraphes quatre et cinq suivants, deux aspects de l'étude de la commande tactile en véhicule :

1. Dans un premier temps, afin de savoir si l'utilisation de la surface tactile est compatible avec la tâche de conduite, il paraît intéressant d'examiner s'il est possible de tracer la commande gestuelle sur la surface tactile sans détourner le regard de la route.
2. Les symboles de commande doivent être choisis de telle sorte qu'ils n'augmentent pas la charge cognitive du conducteur. Une démarche systématique est utilisée pour sélectionner le vocabulaire de gestes.

4. Etude d'une utilisation sans retour visuel de la tablette tactile

4.1. Reconnaissance de caractères

Le module "Reconnaissance" de la Figure 3 repose sur une méthode connexionniste appelée *DVQ* (Dynamic Vector Quantization) (Poirier et Ferrieux, 1991 ; Kamp, 1998a). Parmi les techniques connexionnistes, la carte auto-organisatrice de Kohonen (Kohonen, 1995), la *LVQ* (Kohonen, 1988) et la *DVQ* ont toutes comme point commun d'effectuer un apprentissage compétitif pour construire le réseau. Contrairement à la carte auto-organisatrice qui est destinée à la classification automatique, la *LVQ* et la *DVQ* sont des méthodes où l'apprentissage est de type supervisé. A ce titre, elles sont souvent utilisées en reconnaissance des formes.

Dans un contexte industriel (Renault en l'occurrence), notre but n'est pas de sélectionner "le meilleur" reconnaisseur mais plutôt "un bon" reconnaisseur qui soit rapide et peu gourmand en terme de mémoire. C'est pourquoi, pour évaluer les performances de la méthode *DVQ*, cette dernière a été confrontée à un bon reconnaisseur (Holmström et al., 1997 ; Michie et al., 1994) couramment utilisé en reconnaissance des formes : l'algorithme des *kPPV* ou "k plus proches voisins" (Cover et Hart, 1967). Les résultats issus d'une comparaison effectuée sur une base de données de 5252 lettres majuscules tracées par 101 scripteurs montrent que la *DVQ* offre différents avantages (Kamp et al., 1998b) :

- le taux moyen de reconnaissance s'élève à 93,6 % au lieu de 92,6 % pour les *kPPV* (avec $k=1$),
- le temps de calcul est 100 fois plus faible que celui nécessaire à la méthode de base des *kPPV*.

En dépit des limitations connues de *kPPV*, les résultats montrent clairement que la *DVQ* est une "bonne" méthode de reconnaissance et qu'elle est nettement plus rapide.

4.2. Objectifs de l'étude

La base de données sur laquelle sont testées les méthodes DVQ et kPPV contient des caractères tracés sur la surface tactile dans des conditions bien précises : l'utilisateur ne possède aucun retour visuel de son tracé mais il observe son doigt évoluer sur la surface sensible. Or, dans un véhicule en mouvement, il est vivement conseillé de maintenir son regard sur la route. En pratique, les ergonomes estiment que le délai maximum à ne pas dépasser est de deux secondes (Noy, 1997).

Pour évaluer comment la tablette tactile se comporte face au problème de détournement du regard en roulant, il faut d'abord se poser la question de savoir si, dans un environnement statique (c'est-à-dire à l'arrêt), il est possible d'interagir facilement avec la tablette en regardant droit devant soi. Concrètement, cette section essaie de répondre à deux questions essentielles pour la suite de l'étude :

- est-il possible de tracer les symboles sur la surface tactile sans bénéficier d'aucun retour visuel à l'écran et sans suivre le doigt des yeux (tracé à l'aveugle) ?
- quelle est la robustesse de reconnaissance lorsque, d'une part, l'utilisateur trace les symboles à l'aveugle et, d'autre part, lorsqu'il dispose d'une rétroaction visuelle à l'écran ?

4.3. Expérimentation à l'aveugle





L'expérience se déroule en laboratoire. Chaque sujet qui participe à l'expérimentation est assis à une table. A sa droite ou à sa gauche (droitier ou gaucher) la surface tactile y est fixée à plat. Devant lui à hauteur des yeux, se trouve un écran de 17 pouces. Dans une phase initiale, le sujet se familiarise avec la tablette tactile en écrivant en moyenne une fois l'alphabet des lettres. Lors de l'enregistrement d'un caractère quelconque, le scripteur et/ou l'assistant³ peut suivre le tracé du doigt "en temps réel" dans une fenêtre. Cette fenêtre de taille 256x256 pixels reproduit exactement la forme du tracé puisque le pavé tactile possède précisément une résolution de 256x256 points. Le pixel de l'écran étant à peu près deux fois plus grand (en largeur et en hauteur) que le point de la surface tactile, le caractère à l'écran est agrandi par rapport à sa taille réelle.

Avant chaque enregistrement, le sujet reçoit par écrit les consignes suivantes :

- tracer sur la surface tactile, successivement et uniquement avec le doigt, les lettres majuscules, les chiffres et quatre symboles de commande (Tableau 1),
- la liberté de la forme donnée aux tracés est totale,

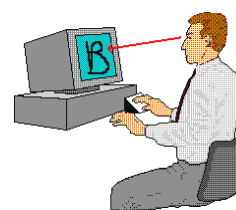
³ L'assistant est la personne qui veille au bon déroulement des enregistrements.

Tableau 1 – Les quatre symboles de commande

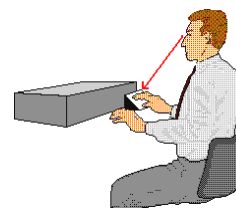
Symbole	Action associée (à titre d'exemple)
	Supprimer l'élément de la liste
	Remonter à la racine de l'arborescence
	Revenir au menu précédent
	Passer au menu suivant

- les tracés sont effectués dans trois conditions différentes :

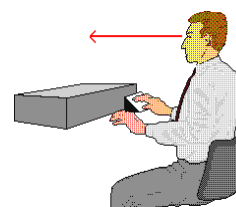
- Condition 1 : Pour chaque caractère⁴ dessiné, il est imposé au sujet d'observer l'évolution de son tracé à l'écran (tracé avec retour visuel), l'écran étant déporté par rapport à la surface tactile.



- Condition 2 : Pour chaque caractère dessiné, il est imposé au sujet de regarder son doigt évoluer sur la surface tactile, l'écran d'ordinateur étant caché (tracé avec retour de la position du doigt).



- Condition 3 : Pour chaque caractère dessiné, il est imposé au sujet de regarder devant lui sans observer ni son doigt, ni l'écran d'ordinateur (tracé à l'aveugle). Dans cette condition, le sujet peut regarder où il pose son doigt sur la surface tactile uniquement au début du tracé de chaque caractère.



- dans chacune des trois conditions d'enregistrement, le sujet trace deux fois chaque alphabet (lettres, chiffres, symboles de commande) pour récolter le maximum de données.

Au total, 101 scripteurs différents ont été enregistrés, récoltant ainsi un corpus de 24240 caractères. Les sujets sont recrutés principalement dans le milieu

⁴ Pour l'expérimentation, un caractère désigne une lettre majuscule, un chiffre ou un symbole de commande.

universitaire et parmi le personnel de Renault. Ils proviennent d'une population composée essentiellement de droitiers (92 %) et de personnes familiarisées avec les outils informatiques, c'est-à-dire : administratifs (12 %), ingénieurs (19 %), enseignants (24 %), étudiants (37 %), autres (8 %).

4.4. Résultats et analyse

Deux types d'analyses peuvent être menées à partir des corpus de caractères. La première est une analyse qualitative des tracés, la seconde est une analyse quantitative (taux de reconnaissance).

4.4.1. Analyse qualitative

La forme des caractères dessinés montre que tous les sujets sont capables de tracer les lettres à l'aveugle : un exemple pour les lettres G, H, I est donné à la Figure 4 pour un même sujet. Bien que le scripteur ne dispose d'aucun retour visuel de son tracé, il est tout à fait capable d'enchaîner sans se tromper les segments qui composent le caractère. Pour la lettre "H", par exemple, le segment horizontal est non seulement correctement placé à mi-hauteur des deux segments verticaux, mais en plus, le raccordement est presque parfait. Le même phénomène est observé pour les deux segments horizontaux de la lettre "I". Cette observation rejoint une hypothèse formulée par certains chercheurs en reconnaissance de l'écriture manuscrite (Plamondon 1989 ; Guerfali 1996). L'hypothèse considère que le tracé d'un symbole déjà appris est le résultat d'un mouvement balistique généré par l'image mentale du symbole et non pas par la rétroaction de la position du doigt (ou du stylo) sur la surface d'écriture.

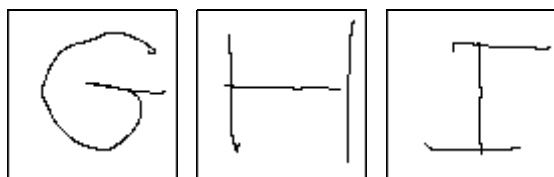


Figure 4 – Tracé des lettres G, H, I par le sujet X dans la condition 3 (à l'aveugle)

A l'inverse, lorsque le sujet regarde le symbole évoluer à l'écran, le trait oscille fortement (Figure 5). La boucle de rétroaction imposée par l'écran empêche le scripteur de produire un mouvement balistique naturel.

La forme des tracés obtenus dans la condition 2 (retour de la position du doigt) est très proche de celle obtenue dans la condition 3 (tracé à l'aveugle) et ne sera donc pas illustrée. Cette constatation est en accord avec l'hypothèse du mouvement balistique naturel : dans la condition 2, le sujet trace le caractère uniquement grâce à l'image mentale du symbole.

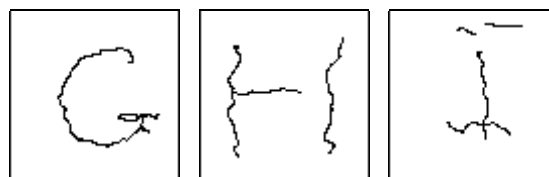


Figure 5 – Tracé des lettres G, H, I par le même sujet X dans la condition 1 (retour visuel à l'écran)

4.4.2. Analyse quantitative

L'analyse quantitative mesure les performances de la reconnaissance lorsque les tracés s'effectuent dans les conditions 1, 2 et 3.

Comme discuté au paragraphe 4.1, la méthode connexionniste de classification utilisée est la DVQ. Les tests de reconnaissance sont effectués sur les lettres majuscules. Dans ce cas, pour chacune des trois conditions d'enregistrement nous disposons d'un nombre total de 5252 lettres tracées par 101 sujets différents. Sur la première moitié des données, on construit le réseau de neurones tandis que sur la seconde on teste la reconnaissance. Les lettres utilisées pour le test sont tracées par des scripteurs tous différents de ceux ayant servi à l'apprentissage. Les taux moyens de reconnaissance obtenus dans chacune des trois conditions sont les suivants :

- condition 1 : 90,6 % ($\sigma = 0,91$),
- condition 2 : 93,8 % ($\sigma = 0,85$),
- condition 3 : **93 %** ($\sigma = 0,93$).

Le meilleur résultat est atteint (93,8 %) lorsque le scripteur regarde son doigt évoluer sur la surface (condition 2). Mais cette valeur n'est pas significativement plus élevée que celle obtenue (93 %) lorsque le scripteur trace les lettres "à l'aveugle" (condition 3). Par contre, le taux moyen de reconnaissance est moins bon (90,6 %) dès que le scripteur doit suivre son tracé "en temps réel" à l'écran (condition 1). Ces résultats sont confirmés par un test de comparaison de moyennes (Saporta, 1990) qui montre que la condition 1 se distingue clairement des conditions 2 et 3 et que les conditions 2 et 3 ne sont pas significativement différentes.

4.4.3. Commentaires

L'ensemble des résultats d'analyse permet d'affirmer qu'en utilisant le doigt comme ustensile d'écriture et la surface tactile comme capteur, l'être humain est tout à fait capable de tracer correctement n'importe quelle lettre majuscule sans bénéficier d'aucun retour visuel. De plus, dans cette condition, le taux de reconnaissance moyen s'élève à 93 %. Ce résultat est très encourageant. Il permet d'envisager l'utilisation de systèmes complexes en véhicule dans la mesure où l'interaction proposée perturbe le moins possible la

tâche de conduite. Bien entendu, le tracé d'un symbole avec le doigt impose de conduire avec une seule main. Cette contrainte devra également être examinée.

5. Génération du vocabulaire de dialogue

Dans cette partie, l'objectif est d'exposer la démarche suivie pour générer et sélectionner les gestes (symboles) de commande qui seront tracés par l'utilisateur sur la surface tactile.

5.1. Analyse des tâches de l'utilisateur

Avant d'émettre des idées quant aux gestes qu'il serait intéressant de reconnaître avec une surface tactile en véhicule, il faut tout d'abord mener une analyse des tâches susceptibles d'être accomplies par le conducteur.

Comme annoncé au paragraphe 2, différentes prestations (systèmes embarqués) sont ou seront disponibles dans les véhicules. Trois d'entre elles sont

particulièrement intéressantes car ces prestations rassemblent quasiment à elles seules l'entièreté des actions (voir Tableau 2) nécessaires à la réalisation de n'importe quelle tâche d'interaction : il s'agit du système de navigation, du téléphone et de la radio. L'analyse de la tâche est effectuée pour chacune des trois prestations.

5.1.1. La navigation

Le système de navigation CARMINAT (Renault, 1995) propose au conducteur un itinéraire routier lorsque ce dernier choisit une destination en accord avec ses préférences personnelles : l'itinéraire calculé peut être le plus rapide en temps ou le plus court en distance. Le système tient compte éventuellement des problèmes de circulation et suggère alors un itinéraire bis. En complément de la navigation, un système d'information est à la disposition de l'utilisateur en ce qui concerne le trafic routier ou les services disponibles tels que les parkings ou les hôtels. Dans ce cas, une carte routière et un curseur apparaissent à l'écran. Le curseur peut

Tableau 2 – Liste des actions possibles

N°	Prestations ⇒	Navigation	Téléphone	Radio
	Actions les plus courantes ↓			
1	Déplacer un curseur dans une liste	Tâches T1, T3, T4	Tâches T1, T3, T4	Tâches T1, T2, T3, T4
2	Déplacer un curseur sur une carte	Tâche T5	Néant	Néant
3	Faire apparaître un sous-menu	Tâche T2	Tâche T2	Néant
4	Sélectionner (valider) un élément dans une liste	Tâches T1, T3, T4	Tâches T1, T3, T4	Tâches T1, T2, T3, T4
5	Rechercher un élément alphanumérique dans une liste	Tâche T4	Tâche T4	Tâche T2
6	Entrer dans une prestation	Tâche T1	Tâche T1	Tâche T1
7	Écrire/rentrer un nouvel élément alphanumérique	Néant	Tâches T5, T7	Néant
8	Enregistrer un nouvel élément	Néant	Tâche T5	Tâche T4
9	Supprimer un élément	Néant	Tâche T6	Néant
10	Annuler/Défaire	Il doit être possible d'exécuter ces tâches quel que soit le contexte		
11	Demander de l'aide			
12	Retourner à la racine			
13	Régler le volume du son			

être déplacé pour sélectionner une zone sur la carte. En relation avec la position du curseur, les occupants choisissent d'être informés sur la circulation routière ou les services. S'ils le désirent, ils peuvent également agrandir ou rétrécir la carte (zoom avant ou arrière).

Pour communiquer avec le système de navigation, le conducteur exécute un certain nombre de tâches parmi les suivantes :

- T1. entrer dans la prestation " navigation ",
- T2. dérouler la liste des sous-menus,
- T3. entrer dans le sous-menu " choisir sa destination " ou " obtenir des informations ",
- T4. choisir dans une liste de noms, la destination, ou le type d'information désiré (parkings, hôtels),
- T5. déplacer le curseur sur la carte (suivant 2 directions),
- T6. agrandir ou rétrécir la carte.

5.1.2. Le téléphone

Outre le fait qu'il soit possible de composer un numéro, les téléphones actuels permettent de gérer un répertoire. De cette manière il suffit de sélectionner le nom du correspondant sans devoir former son numéro. L'utilisateur doit donc pouvoir réaliser les tâches relatives à la gestion d'une liste d'éléments : sélectionner, ajouter, supprimer un élément.

Le conducteur qui interagit avec le téléphone, accomplit une ou plusieurs des tâches suivantes :

- T1. entrer dans la prestation " téléphone ",
- T2. dérouler la liste des sous-menus,
- T3. entrer dans le sous-menu " ajouter un élément " ou " supprimer un élément ",
- T4. choisir dans une liste de noms, un correspondant téléphonique,
- T5. écrire un nouveau correspondant dans la liste des noms,
- T6. supprimer un correspondant de la liste des noms,
- T7. composer un numéro de téléphone quelconque.

5.1.3. La radio

C'est l'équipement le plus familier. Pour cette prestation, on se limite aux manipulations les plus courantes de la radio sans considérer les réglages fins du son (graves, aiguës, balance) rarement effectués.

Dans ce cas, les tâches pouvant être réalisées par le sujet sont les suivantes :

- T1. entrer dans la prestation " radio ",
- T2. choisir dans une liste de stations émettrices préenregistrées, la station désirée,
- T3. entrer dans le sous-menu " syntonisation " (méthode de réglage des récepteurs radio),
- T4. effectuer une recherche en fréquence pour sélectionner une station émettrice, ou pour enregistrer une nouvelle station.

5.1.4. Tableau récapitulatif

Sans prétendre à l'exhaustivité, l'ensemble des actions les plus courantes pour les trois prestations navigation, téléphone et radio sont rassemblées dans le Tableau 2. La numérotation des tâches reprises dans ce tableau correspond à la numérotation adoptée aux paragraphes 5.1.1, 5.1.2, 5.1.3.

Le bas du tableau (actions 10 à 13) regroupe les actions que l'utilisateur peut accomplir quel que soit le contexte. En effet, certains symboles de commande doivent être reconnus par le système à tous moments :

- " annuler " est indispensable pour corriger les erreurs ou revenir en arrière,
- " régler le volume du son " doit être exécutable immédiatement,
- " demander de l'aide " et " retourner à la racine " ne peuvent pas dépendre de la prestation.

5.2. Conception participative

Après avoir identifié les tâches susceptibles d'être exécutées par l'utilisateur, l'étape suivante consiste à générer les symboles dédiés à l'accomplissement de chaque tâche.

La méthode utilisée est celle de la conception participative (" brainstorming ") (Byrne et Barlow, 1993). L'objectif est de rassembler un groupe de personnes afin d'étudier ensemble la création d'un système ou produit novateur. En ce qui concerne notre étude, il s'agit du système tactile. Plusieurs réunions ont été nécessaires à l'accomplissement de ce travail. Les participants étaient au nombre de quatre : un ingénieur de recherche et un designer, tous deux de Renault, et deux chercheurs universitaires. À chaque réunion, la démarche adoptée était la suivante : pour chacune des actions du Tableau 2 (numérotées de 1 à 13), s'est posée la question de savoir comment l'utilisateur peut le traduire à l'aide d'un symbole tracé sur la surface tactile. Dans certains cas exceptionnels, un bouton est proposé à la place d'un symbole. En effet, pour certaines tâches qui doivent être réalisées rapidement et sûrement (le contrôle du volume sonore par exemple), un bouton peut être préférable.

Parmi l'ensemble des symboles (que nous ne montrerons pas afin de ne pas surcharger l'exposé) créés suivant la technique du " brainstorming ", une sélection importante est réalisée sur la base de différents critères :

- minimiser le nombre de boutons,
- réduire au minimum les temps de manipulation et donc le nombre d'étapes dans la séquence des opérations pour réaliser une action donnée,

Tableau 3 – Liste des gestes sélectionnés

Actions ↓	Symbole N°	Tracé	Utilisation – Commentaires			
❶ Déplacer un curseur dans une liste	1		Appui sur une des 4 zones et dans la direction désirée. L'appui long accélère le curseur.			
	2 (♣)		Tracé d'un trait dans une des 4 directions possibles, n'importe où sur la surface. L'appui long accélère le curseur.			
❷ Déplacer un curseur sur une carte	2 (♣)		Reprise du symbole N°2 et ajout de la possibilité d'un déplacement en diagonale.			
	3		La position du doigt sur la surface tactile circulaire donne la direction de déplacement du curseur (α) et la vitesse de déplacement (d).			
	4	<table border="1" style="display: inline-table; vertical-align: middle;"> <tr> <td>Surf. Tact.</td> <td>Carte</td> </tr> <tr> <td>• (x,y)</td> <td>• (x,y)</td> </tr> </table>	Surf. Tact.	Carte	• (x,y)	• (x,y)
Surf. Tact.	Carte					
• (x,y)	• (x,y)					
❸ Faire apparaître un sous-menu	5 (♣)	<table style="display: inline-table; vertical-align: middle;"> <tr> <td>(1) </td> <td>(2) </td> </tr> </table>	(1)	(2)	Le mouvement circulaire du doigt (1) entraîne l'affichage d'un menu "camembert" (2) à l'écran.	
(1)	(2)					
❹ Sélectionner (valider) un élément	6 (♣)		Appui court n'importe où sur la surface tactile.			
❺ Rechercher un élément alphanumérique dans une liste	7 (♣)		Reconnaissance successive du tracé des lettres/chiffres qui composent l'élément alphanumérique recherché dans la liste.			
	8 (♣)		Si la liste est énumérée, la rubrique est sélectionnée en traçant le chiffre correspondant.			
❻ Entrer dans une prestation	9 (♣)		Par reconnaissance du tracé de l'initiale (N pour navigation, T pour téléphone, R pour radio), la prestation est directement sélectionnée.			
	10	<table style="display: inline-table; vertical-align: middle;"> <tr> <td>(1) </td> <td>(2) </td> </tr> </table>	(1)	(2)	Par reconnaissance d'un symbole donné (1), la barre du menu principal des prestations (2) est affichée à l'écran.	
(1)	(2)					
❼ Ecrire/rentrer un nouvel élément alphanumérique	11 (♣)		Reconnaissance successive du tracé des lettres/chiffres qui composent l'élément alphanumérique à écrire.			
❽ Enregistrer un nouvel élément	12 (♣)		Tracé de la lettre majuscule " E ", faisant allusion au mot " Enregistrer ".			
	13 (♣)		Tracé de la lettre majuscule " V ", faisant allusion au mot " Valider " ou " OK ". Le symbole 13 est tracé plus rapidement que le symbole 12 qui demande 2 à 3 levés du doigt.			
❾ Supprimer un élément	14 (♣)		Tracé du geste " rature ". Geste mnémotechnique et connu dans la littérature (Rubine, 1991).			
❿ Annuler/Défaire	Bouton N°15		L'action " annuler " ne peut pas se confondre avec une autre action : un bouton est donc choisi.			
⓫ Demander de l'aide	16 (♣)		Tracé du symbole " point d'interrogation ".			
⓬ Retourner à la racine	17 (♣)		Le tracé du symbole fait immédiatement apparaître la barre principale des menus. Il doit être identique au symbole N°10 car l'action est identique.			
⓭ Régler le volume du son	Boutons N°18		volume+ / volume- Les boutons sont indispensables pour permettre un contrôle rapide et sûr du volume sonore.			

- choisir des symboles qui soient discriminants entre eux pour faciliter la reconnaissance,
- choisir des symboles mnémotechniques sachant qu'un être humain mémorise un maximum de 5 à 7 éléments différents sans se tromper (Miller, 1956).

Le Tableau 3 montre pour chaque action les symboles sélectionnés. Dans bien des cas, plusieurs gestes alternatifs sont volontairement conservés pour une même action car il n'est pas possible de prendre de décision sans tester réellement le symbole à ce stade de l'étude (actions 1, 2, 5, 6, 8). Les gestes qui ne nécessitent pas de détourner le regard pour observer le tracé du doigt sur la surface tactile sont annotés du symbole "♣". Les actions associées à ces gestes devraient être compatibles avec la tâche de conduite.

Conclusion

Après avoir montré les avantages de l'interaction gestuelle avec les systèmes embarqués, nous avons accompli une première étape dans l'étude de la commande tactile en véhicule, à savoir :

1. mener une évaluation ergonomique qui aboutit à un résultat encourageant concernant l'adéquation de la commande gestuelle avec la conduite automobile. Lors de cette évaluation, 101 scripteurs ont été capables de tracer des lettres majuscules à l'aveugle en laboratoire. Dans cette condition, le taux de reconnaissance s'élève à 93 %.
2. sélectionner un vocabulaire de dialogue suivant la méthode de la conception participative. Les symboles choisis doivent également vérifier trois critères importants par rapport au contexte d'utilisation en véhicule : temps minimum pour accomplir une action, symboles discriminants pour faciliter la reconnaissance et préférence pour des symboles mnémotechniques.

Rien n'indique que lorsque le sujet sera dans une double tâche de conduite et d'interaction avec un système embarqué, il pourra atteindre les performances mentionnées dans l'expérimentation en laboratoire. Dès lors, dans une étape ultérieure, il sera intéressant d'étudier le comportement de la surface tactile dans des conditions réelles c'est-à-dire à bord d'un véhicule en mouvement. Sur base de mesures liées au contrôle de la trajectoire et de la vitesse, la perturbation de la tâche de conduite par l'usage des dispositifs de commande pourra être évaluée.

Etant donné le succès grandissant de la modalité parole adaptée à une utilisation en véhicule (VODIS, 1996), il est prévu, à plus long terme, d'examiner les technologies vocale et gestuelle dans une communication multimodale. Pour cela, il faudra

considérer leurs complémentarités dans le contexte d'une interaction en véhicule.

Références bibliographiques

[Bellik, 1995] Y. Bellik, (1995), Interfaces multimodales : concepts, modèles et architectures. Thèse de l'Université de Paris-XI.

[Byrne et Barlow, 1993] J.G. Byrne, T. Barlow, (1993), Structured Brainstorming : A Method for Collecting User Requirements. Proceedings of the 37th Annual Meeting Human Factors and Ergonomics Society, Vol. 1, 427-431.

[Catling, 1994] I. Catling, (1994), Advanced Technology for Road Transport : IVHS and ATT. I. Catling (Eds), Artech House, 1-376.

[Coleman et al., 1991] M.F. Coleman, B.A. Loring, M.E. Wiklund, (1991), User Performance on Typing Tasks Involving Reduced-Size, Touch Screen Keyboards. VNIS'91, Proc. of the Vehicle Navigation and Information Systems Conference, New York, NJ, 534-549.

[Cover et Hart, 1967] T. Cover, P.E. Hart, (1967), Nearest Neighbour Pattern Classification. *IEEE Transactions Information Theory*, 13: 21-27.

[Franzen, 1991] S. Franzen, (1991), Conventional Controls. Drive Project V1041, Analysis of input and output devices for in-car use, 7.1-7.30.

[Guerfali, 1996] W. Guerfali, (1996), Modèle delta lognormal vectoriel pour l'analyse du mouvement et la génération de l'écriture manuscrite. Thèse de l'Université de Montréal, École Polytechnique de Montréal.

[Holmström et al., 1997] L. Holmström, P. Koistinen, J. Laaksonen, E. Oja, (1997), Neural and Statistical Classifiers – Taxonomy and Two Case Studies. *IEEE Transactions on Neural Networks*, Vol.8, N°1, 5-17.

[Kalisvaart et Hartemann, 1991] B. Kalisvaart, F. Hartemann, (1991), Touchscreens. Drive Project V1041, Analysis of input and output devices for in-car use, 6.1-6.14.

[Kamp, 1998a] J.F. Kamp, (1998), Interaction Homme-Machine à bord des véhicules automobiles : Etude des modalités et des dispositifs d'interaction. Thèse de l'Ecole Nationale Supérieure des Télécommunications, Paris, 1-223.

[Kamp et al., 1998b] J.F. Kamp, F. Poirier, P. Doignon, (1998), Interaction with in-vehicle systems. A neural network approach to recognize symbols drawn with the finger on a touchpad. 15th International Congress on Cybernetics, Namur, Belg., 249-254.

[Kohonen, 1988] T. Kohonen, (1988), Self-Organization and Associative Memory, Springer Verlag, Berlin, RFA.

[Kohonen, 1995] T. Kohonen, (1995), Self-Organization Maps, New York, Springer Verlag, Berlin, RFA.

[Leiser, 1991a] R.G. Leiser, (1991), DRIVE Project V1041, Analysis of input and output devices for in-car use. B. Leiser

and D. Carr (Eds), Traffic Research Centre, University of Groningen, NL.

[Leiser, 1991b] R.G. Leiser, (1991), Speech Output Devices. Drive Project V1041, Analysis of input and output devices for in-car use, 2.1-2.11.

[Leiser, 1993] R.G. Leiser, (1993), Driver-vehicle interface : dialogue design for voice input. *Driving Future Vehicules*, Andrew M. Parkes and Stig Franzen.

[Michie et al., 1994] D. Michie, D.J. Spiegelhalter, C.C. Taylor, (1994), Machine Learning, Neural, and Statistical Classification. C.C. Taylor (Eds), London : Ellis Horwood Ltd.

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

[Neboit, 1994] M. Neboit, (1994), Perception, anticipation et conduite automobile. *Le travail humain*, n°37, 53-72.

[Néel et al., 1996] F. Néel, G. Chollet, L. Lamel, W. Minker, A. Constantinescu, (1996), Reconnaissance et Compréhension de la Parole. In H. Méloni (Eds) *Évaluation et Applications*. AUPELF, 331-367.

[Noy, 1997] Y.I. Noy, (1997), Ergonomics and Safety of Intelligent Driver Interfaces. Y.I. Noy (Eds), Lawrence Erlbaum Associates, 1-432.

[Ouadou et al., 1993] K. Ouadou, B.T. David, G. Labiale, (1993), Interface intelligente dans le poste de conduite automobile et implications ergonomiques. *Interface des mondes réels et virtuels*, INFORMATIQUE 93, Montpellier, France.

[Plamondon, 1989] R. Plamondon, (1989), An evaluation of motor models of handwriting. *IEEE Transactions on Systems, Man and Cybernetics*, vol.19, n°5, 1060-1072.

[Pleczon et al., 1994] P. Pleczon, S. Chalard, M. de Saint Blanchard, (1994), Interface Homme-machine multimodale pour un copilote intelligent d'aide à la conduite. Quatrième colloque Ergo'IA, 108-118, Biarritz, France.

[Poirier et Ferrieux, 1991] F. Poirier, A. Ferrieux, (1991), DVQ : Dynamic Vector Quantization - An incremental LVQ. Proc. ICANN'91, Helsinki, Finlande.

[Price, 1995] P. Price, (1995), Spoken Language Understanding. in (coll.) R. Cole et V. Zue (Eds.) *State of the Art in Human Language Technologies*, chapter 1, *Spoken Language Input*, 49-56.

[Renault, 1995] (1995), PROMETHEUS : fluidité et sécurité routière, *Bulletin Technique*, Renault, 74, 2-33.

[Rubine, 1991] D. Rubine, (1991), The automatic recognition of gestures. PhD thesis, School of Computer Sciences, Carnegie Mellon University.

[Saporta, 1990] G. Saporta, (1990), Probabilités, analyse des données et statistique, Paris : Technip.

[Sears, 1991] A. Sears, (1991), Improving touchscreen keyboards : design issues and a comparison with other devices. *Interacting with Computers*, 3 (3), 253-269.

[Shneiderman, 1998] B. Shneiderman, (1998), Designing the user interface : strategies for effective human-computer interaction. (2nd ed.). Reading, MA : Addison-Wesley.

[Smolders et al, 1994] J. Smolders, R. Diller, D. van Compernelle, (1994), Noise Robust Speech recognition Makes In-car Navigation Safe and Affordable. ICSLP, Yokohama, Japon.

[VODIS, 1996] VODIS, (1996), Voice Operated Driver Information System. Projet Européen, Annual Report. disponible sur la Toile : <http://werner.ira.uka.de/VODIS/>