

Logiciels d'aide
à la communication
pour les sourds
(DATHA)

Objectif de Datha

- Perception de la parole par les sourds, faciliter la communication orale en temps réel et sans contrainte avec les entendants
- Aide à l'intégration :
 - Scolaire
 - Professionnelle
 - Sociale

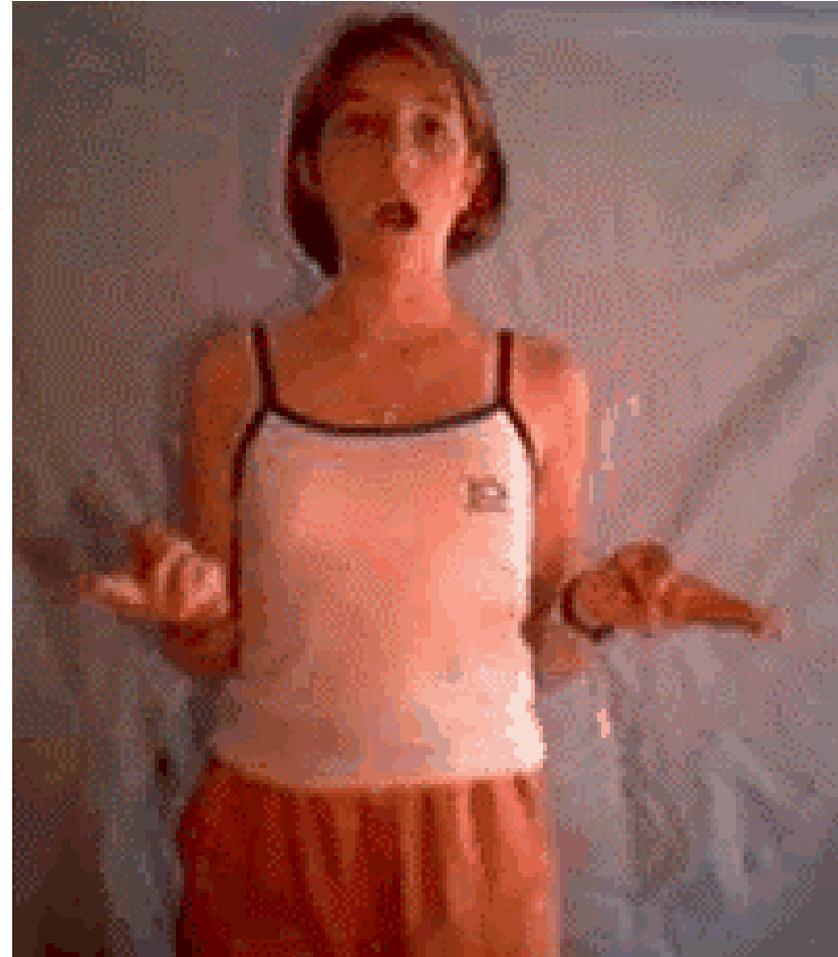
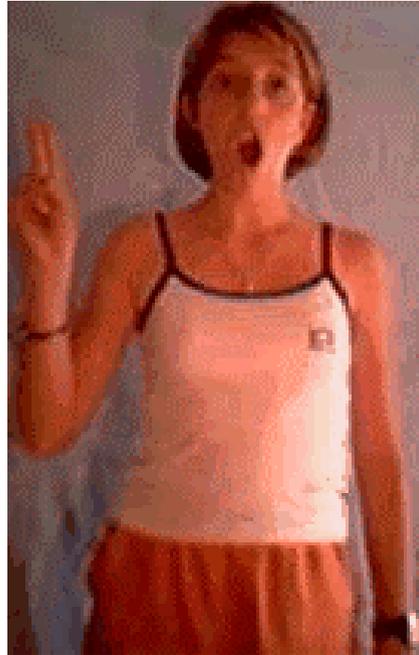
La population sourde

- Evaluée à 7% de la population (4 millions de sourds)
- Environ 1 320 000 sourds moyens
- Environ 500 000 sourds sévères et profonds
- 30 000 enfants sourds profonds
- 420 000 enfants malentendants
- 1 enfant sur 1000 devient sourd avant 18 mois
- 2 enfants sur 1000 deviennent sourds avant 14 ans
- Environ 100 000 pratiquent la LSF

Communication existante

- Langue des Signes
- Lecture labiale
- Langage Parlé Complété
- Sous-titrage

LSF, BSL, ASL



vidéo

La LSF est une langue

- Différente du français
- Avec sa syntaxe et sa grammaire

Un signe est un mouvement = séquence de « visèmes »

Vidéo

Rôle de la lecture labiale dans la perception de la parole

La lecture labiale ne permet pas la perception de tous les contrastes phonologiques: 36 sons correspondent à 12 images labiales

- Phonèmes différents
 - par nasalité /p/versus/m/
 - par voisement /b/versus/p/
 - ont des images identiques sur les lèvres
- Phonèmes invisibles sur les lèvres (k,g,r)
- Phonèmes peu différenciés (e,e,i)

Communication LPC

- LPC : aide à la lecture labiale qui a pour objectif l'acquisition et la transmission de la langue française
- Association LPC + lecture labiale
 - Codage que l'enfant sourd doit acquérir
 - Bénéfiques pour
 - réception du message
 - appropriation langue orale (puis écrite)
 - exercer une suppléance mentale

Au quotidien

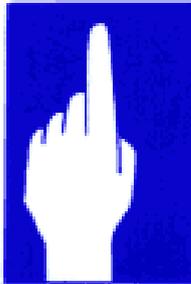
Vidéo : Anne la-
Bégué en live



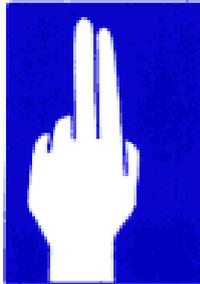
Vidéo

LPC

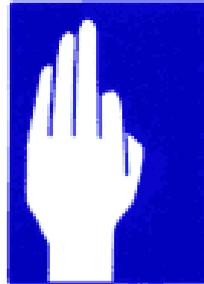
Le Langage Parlé Complété est un code de désambiguïsation labiale : ce n'est pas une langue. La langue est le français.



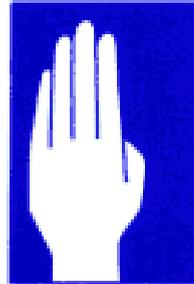
d (dos)
p (par)
3 (joue)



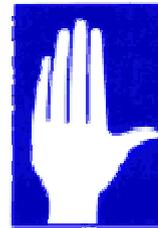
k (car)
v (va)
z (base)



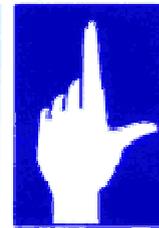
s (sel)
R (rat)



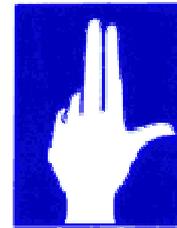
b (bar)
n (non)
y (lui)



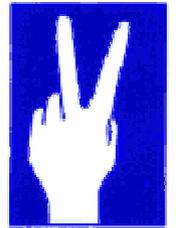
t (toi)
m (mare)
f (fa)
et toute
voyelle non
précédée d'une
consonne
(âge)



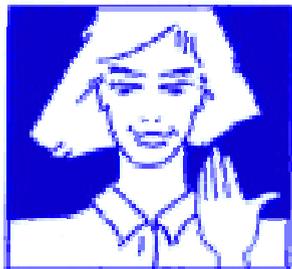
l (la)
ç (chat)
ñ (vigne)
w (ou)



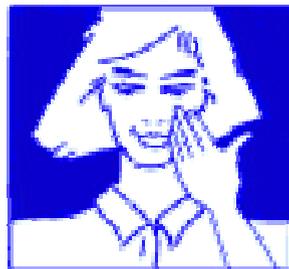
g (gare)



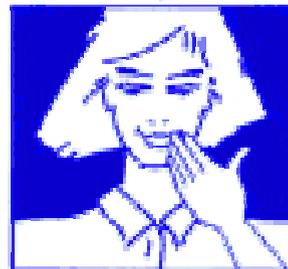
j (fille)
ŋ (camping)



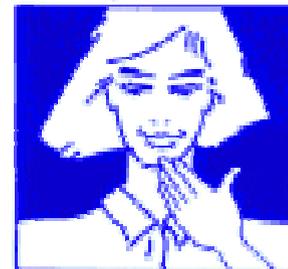
Position côté.
a (ma)
o (maux)
œ (teuf-teuf)
et toute consonne
suivie d'un e muet



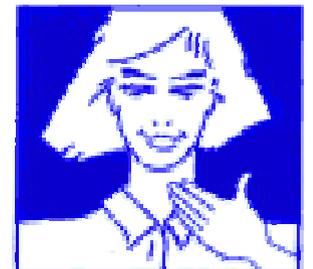
Position pommette.
ë (main)
ø (feu)



Position bouche.
i (mi)
u (ton)
o (man)



Position menton.
E (mais)
u (mou)
ɔ (fort)

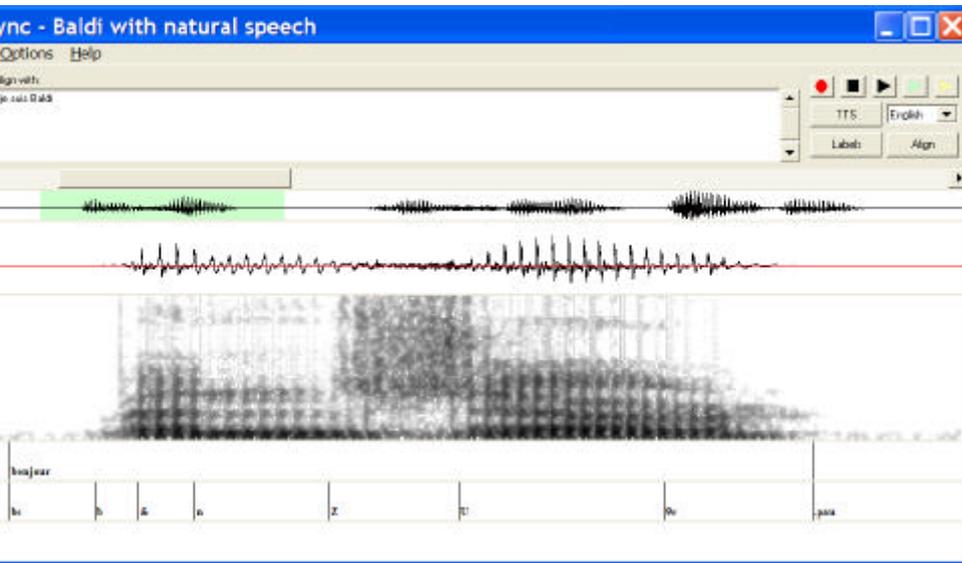


Position gorge.
œ (un)
y (tu)
e (fée)

Position doigts, syllabe position main, consonne

Reconnaissance vocale : français ou phonèmes ?

1. Signal sonore



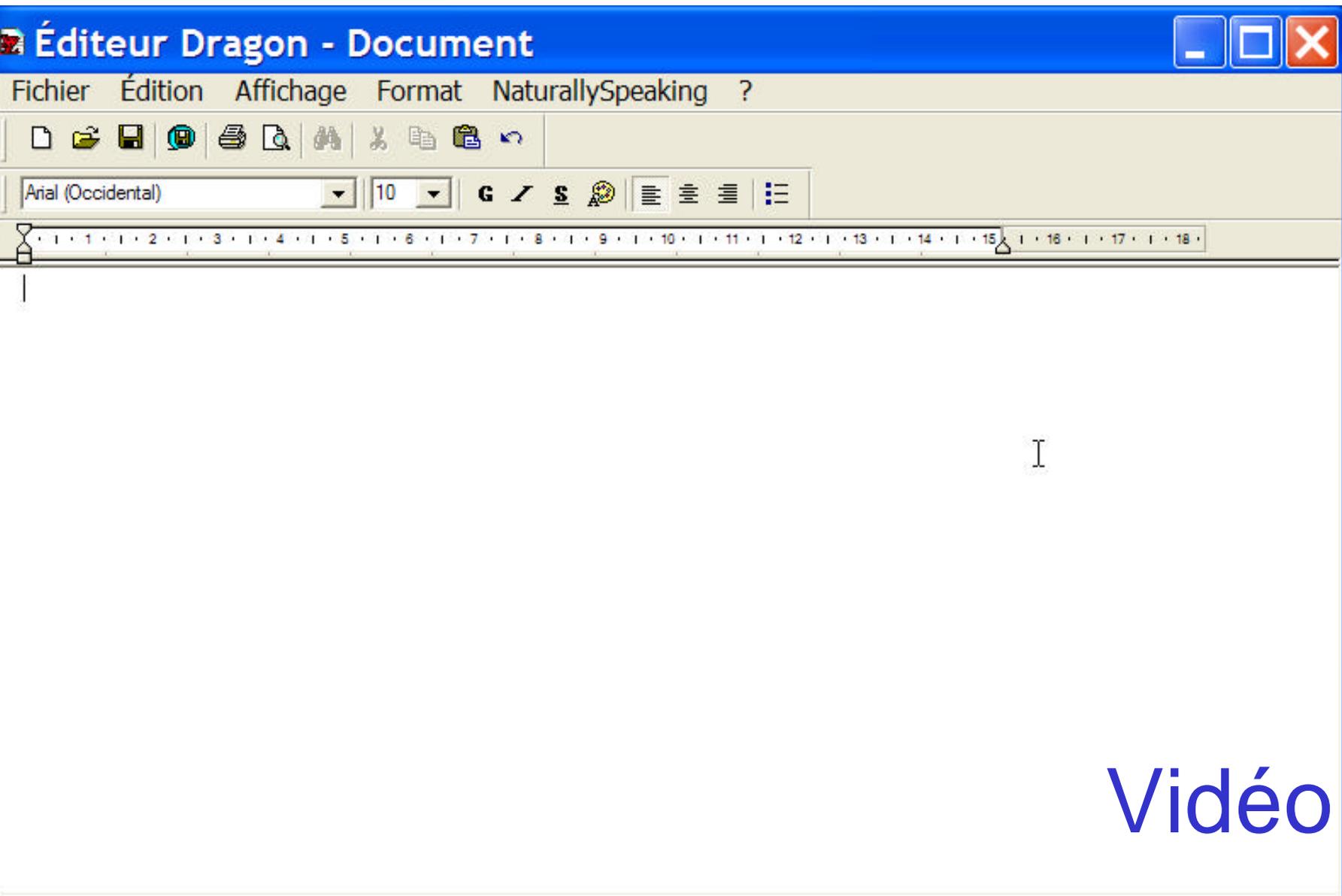
2. Reconnaissance de phonèmes :

« b 60 o~ 60 Z 80 u 190 R 120 Z
70 @ 60 m 50 a 70 p 70 E
40 l 50 a 90 n 80 k 80 a 80
R 90 O 80 l 100 »

3. Reconnaissance du français :

« Bonjour, je m'appelle Anne-Carole »

Dictée Vocale



Difficultés de l'automatisation LSF et sous-titrage

Traduction automatique LSF vers français

- Reconnaissance des signes
- Traduction

Traduction automatique français vers LSF

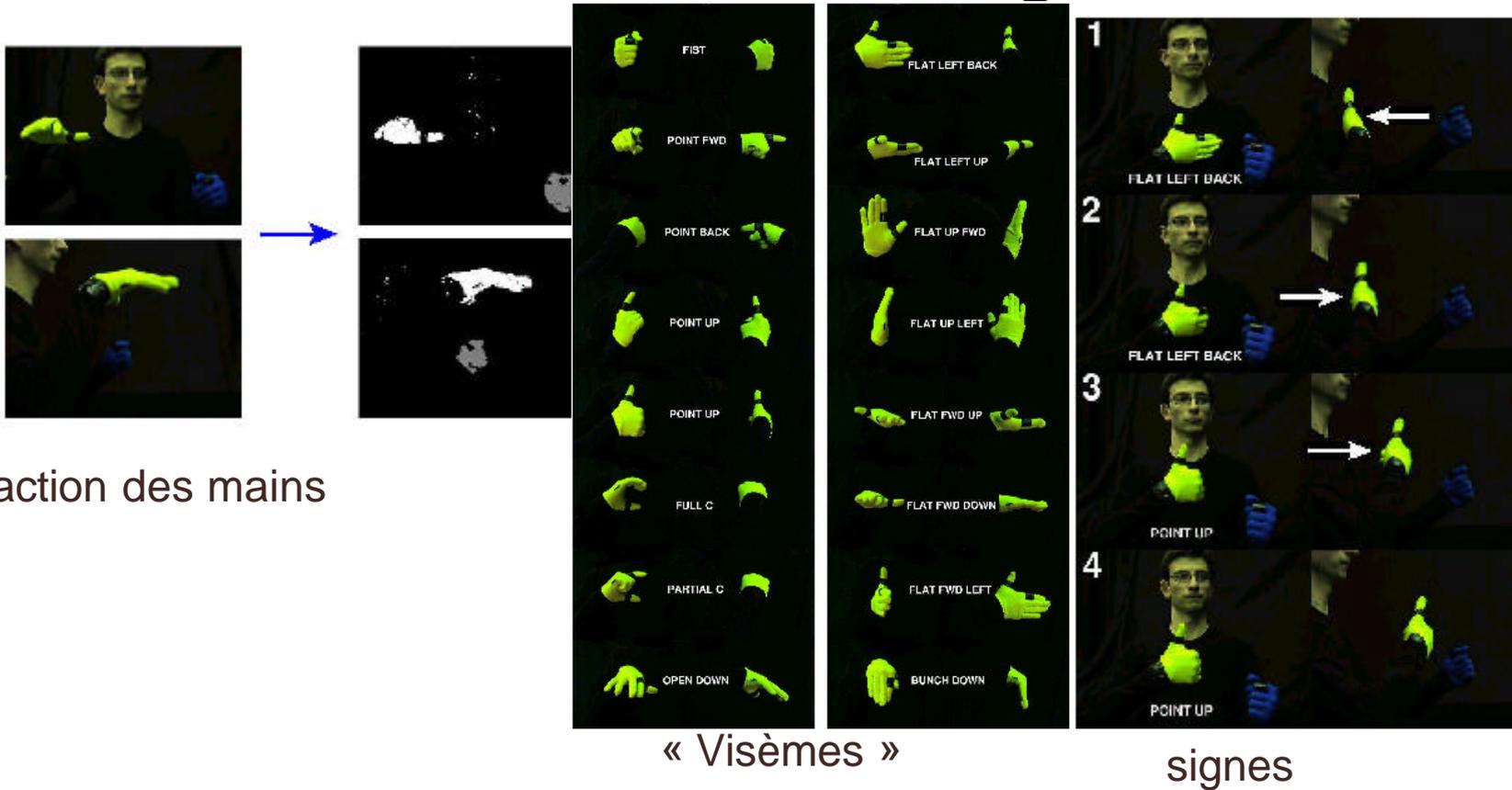
- Reconnaissance du français
- Traduction puis animation 3D

Problème de recherche: INRIA, MIT, Oxford U, INTEL...

Sous-titrage automatique

Reconnaissance du français parlé multi-locuteur

Exemple: LSF->français, reconnaissance des signes



Reste ensuite à traduire en français...

Datha

Association pour le développement d'aides
techniques pour les personnes handicapées

Equipe

- **Anne Lau-Bégué** : ICN, Manager première ligne (RH) EDG/GDF. Administrateur ALPC.
- **Jérémy Boroy** : Diplômé Science-Po Paris. Chargé de communication. Sourd (LPC).
- **Jacques Feldmar** : Docteur INRIA/Polytechnique. Co-fondateur Mirada et LTU. Consultant.
- **Johanna Petit** : Deux maîtrise de Lettres. Responsable association « un Autre Regard ». Sourde (LSF)
- **Pierre Lamothe** : Ingénieur informaticien, EDF, Sourd.
- **Sylvain Beaujouan** : Graphiste sur ordinateur, animateur 3D. Consultant.
- **Frédéric Devernay** : Ingénieur Polytechnique. Docteur INRIA/Polytechnique. Chercheur INRIA.
- **Maryvonne Chéroux (Graphiste), Rémy Zimmerman (Technicien SNCF), Pascal Le Bigot (Médecin anesthésiste) et autres personnes sourdes, parents ou amis...**
- **Support d'institutions, d'associations et de professionnels** : (CNEFEI, ALPC, AFIDEO, un Autre Regard)...

Le projet LPCAO (Datha): Automatisation du LPC (Speech to visual speech)

Cour terme :

Combinaison

- reconnaissance de phonèmes multi-locuteurs
- animation faciale + main

Plus long terme :

- Intégration reconnaissance de la parole (mono-locuteur) pour sous-titrage et sauvegarde textuelle en différé
- Vision par ordinateur pour adaptation de la tête parlante

Différentes combinaisons



Lecture
labiale



Tête
parlante
+
désambig
uisation

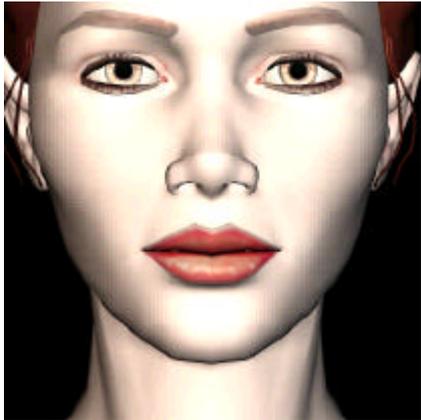


Vidéo +
désambi-
guisation



Sous-titrage
phonétique
ou français
(anglais...)
avec ou
sans main

Un exemple de Datha



Vidéo



Vidéo

Quatre Logiciels Datha

- Apprentissage du français (parlé complété)
- Salle de classe
- Téléphone
- Réunion

Version PC puis PDA/téléphone mobile (GPRS)

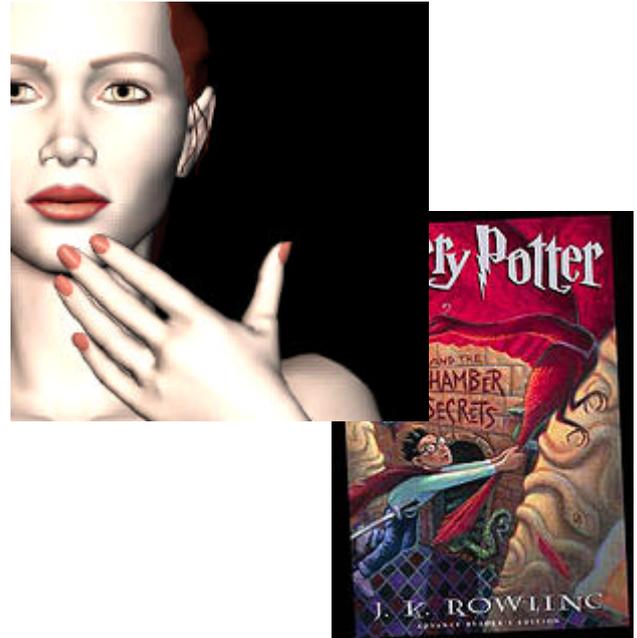
Notre mission :

- Intégration
- Validation
- Promotion

Récits (Texte vers Tête seulement)



Sous-titrage de film



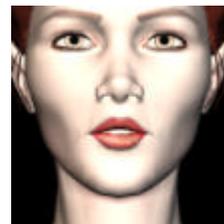
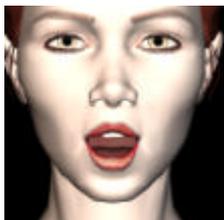
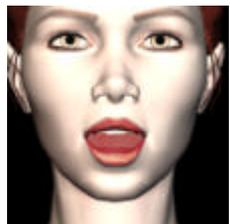
Récit de contes pour enfants

Sous-titrage

Matériel : PC, incrustation TV

Cible : Enfants, adultes

Lecture Labiale



« a »

« a~ »

« p »

« s »

« v »

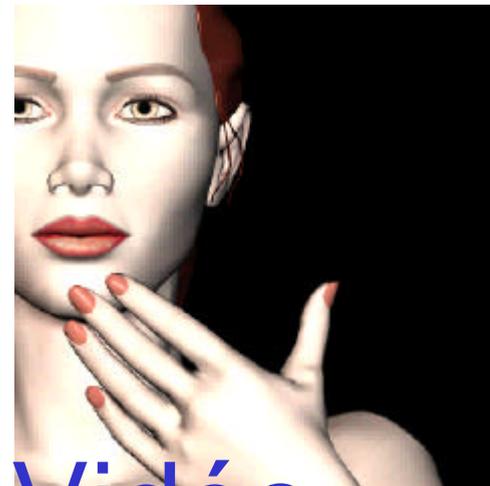
Exemple de visualisation des phonèmes : « a », « a~ », « p », « s », « v »

Exemple de visualisation d'une phrase en lecture labiale

Apprentissage du LPC et de la reconnaissance labiale

Matériel : PC

Cible : Enfants, Parents



Vidéo

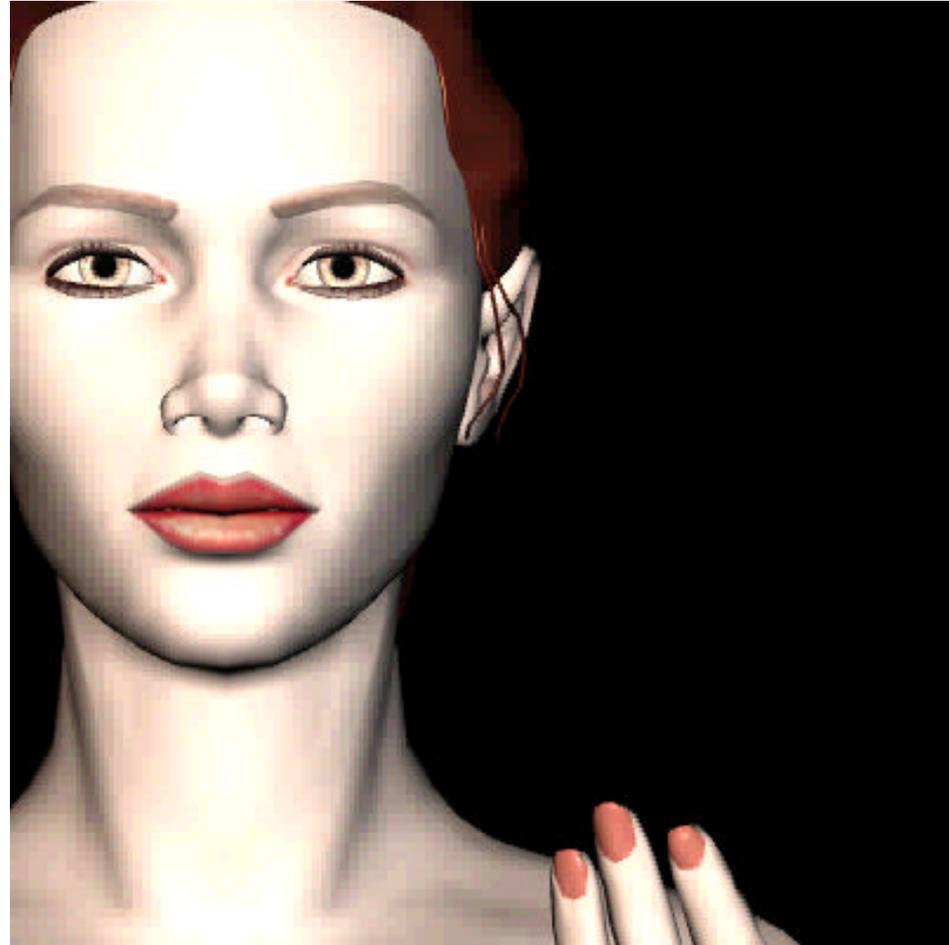
En cours : Apprentissages, Récits

Texte -> LPC

Matériel : PC, incrustation TV

Cible : Enfants, adultes

Vidéo



Vidéo : Sylvain Beaujouan

A l'Ecole ou à l'Université :
codeur automatique avec de la



Téléphone



Vidéo

Traduction direct

Matériel : PC, PDA, mobile

Cible : Enfants, étudiants, adultes



Réunions



Lecture labiale difficile

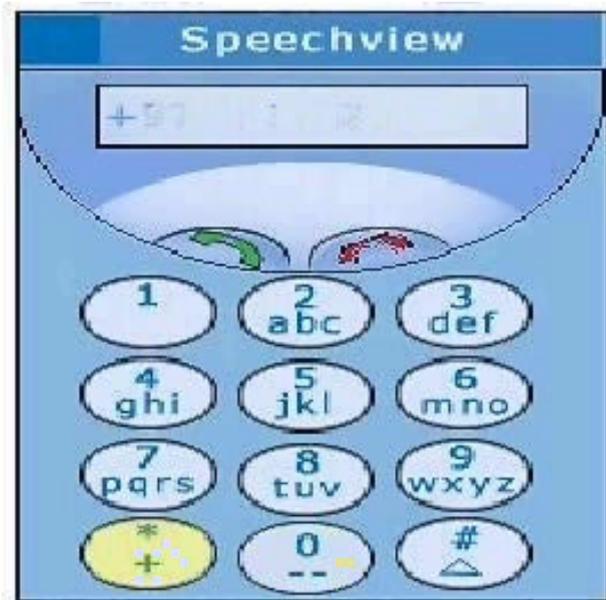
Repérage difficile du locuteur

Problème de la prise de note

Logiciels commerciaux

- Reconnaissance vocale (IBM, Dragon, Philips...)
- Reconnaissance de Phonèmes (Telisma, Babel...)
- SpeechView
- Animation faciale (Poser, FaceAD, SyncMagic...)
- Avatars (la Cantoche. EveMatic. Hantek...)

SpeechView : téléphone

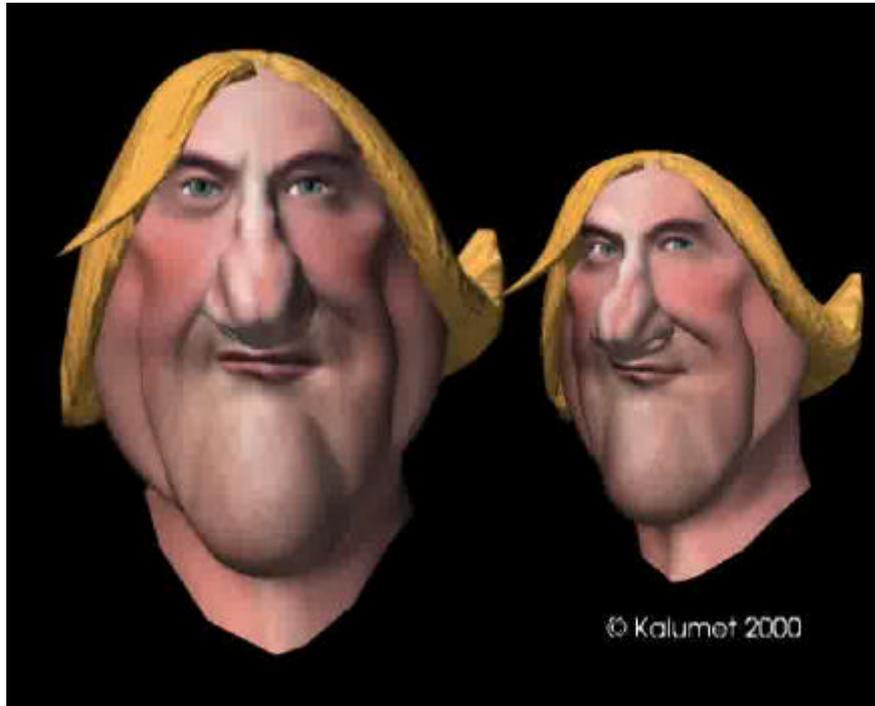


- Logiciel du côté de la personne sourde.
- Pas de contrainte pour l'interlocuteur.
- En 2D seulement, et en anglais...

Vidéo: [speechview courtesy](#)

Vidéo

LipSinc pour la télé ou le cinéma



Vidéo

Autres projets

Prototypes uniquement

- IBM, Lipcom : dans la salle de classe
- Synface (<http://www.speech.kth.se/teleface/>)
- Baldi (<http://cslu.cse.ogi.edu/toolkit/>)
- Voice (http://voice.jrc.it/home_f.htm)
- Projet Artus (ICP, Arte, Thales...)
(<http://www.icp.inpg.fr/ICP/equipes/synthese/avtts/clone.fr.html>, http://www.telecom.gouv.fr/rnrt/projets/res_01_37.htm)

Logiciels universitaires ou Open Source :

- Reconnaissance de phonème ou vocale:
 - Microsoft, Intel, INRIA Lorraine (LORIA), Sphinx
 - LIMSI (CNRS), MIT, CSLU...
- Animation faciale :
 - ICP (CNRS)
 - UCSC, Baldi
 - ATT...

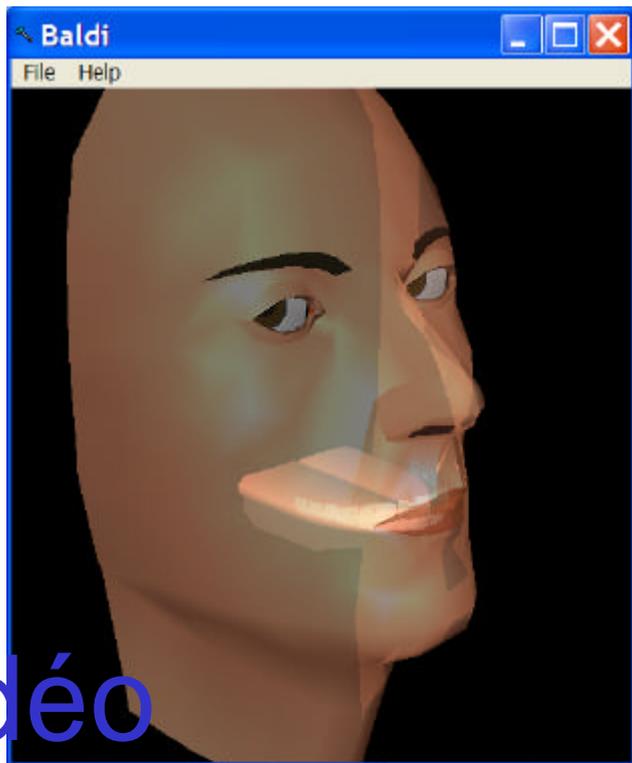
Viavoice/Lipcom/NetScribe

Sur Titrage en français + phonétique. IBM. Projet gelé.



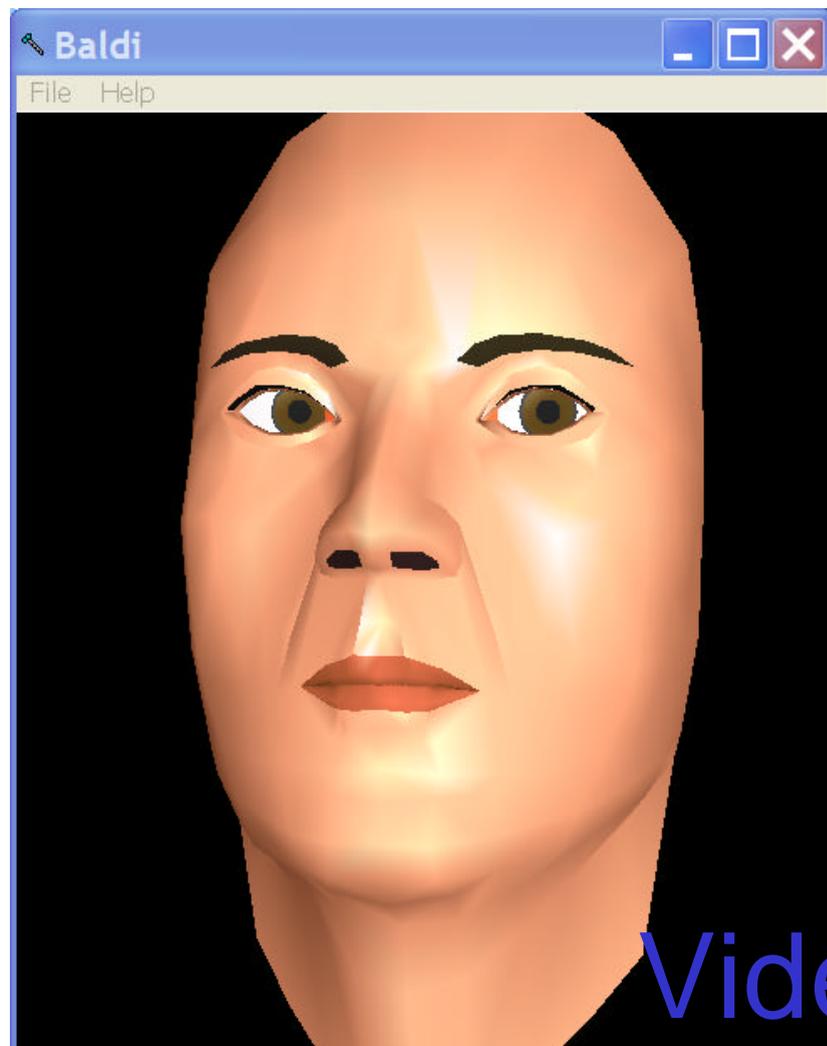
En blanc, le français, **en bleu, le phonétique**, en fonction du score

Baldi



Vidéo

“Hi my name is Baldi. I do not speak French but I might learn. I do not know cued speech either, but you might give me a hand...”



Vidéo

Projet Européen SYNFACE

Vidéo : <http://www.speech.kth.se/teleface/>



La tête
est
animée
par la
voix

Vidéo

Travail sur la coarticulation (Datha/Education Nationale/INRIA)



Vidéo

Mise en œuvre du projet

Moyens nécessaires (sous traitance)

- Logiciel enseignement (LPC, lecture Labiale) :
45KE
- Technologie Texte-to-Tête:
30 KE
- Codeur automatique (Datha+ INRIA):
200KE
- Logiciel Téléphone et Réunions :
en cours d'évaluation (AGEFIPH, FT,
EDF ?)

Ressources

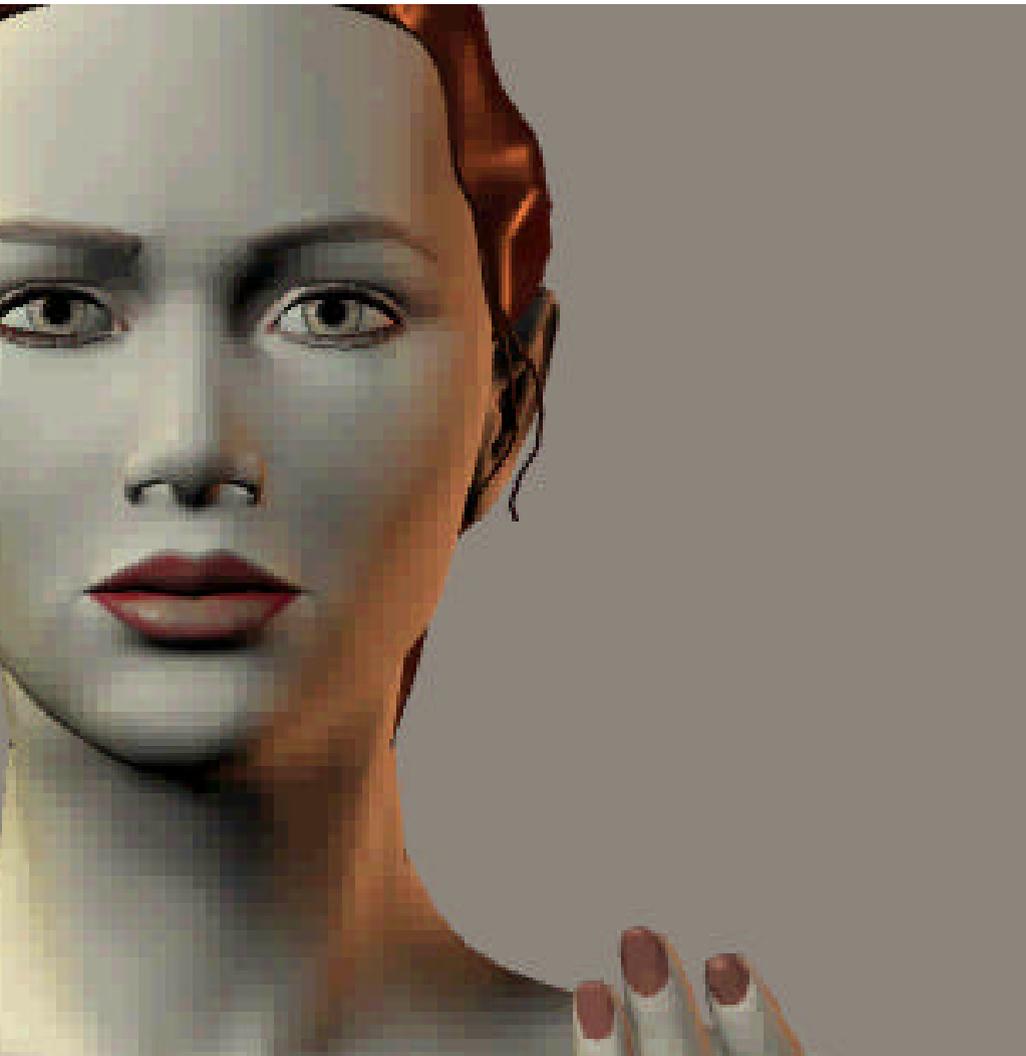
- Coordinateur
- Ergonome et lien utilisateur
- Un ingénieur « mono-poste mono-tête »
- Un ingénieur intégration client-serveur, multi-tête
- Achat de logiciels commerciaux et partenariats

Réseau Transfert de Technologie ?

- Le besoin de Datha est standard :
transfert de technologie vers les utilisateurs (sourds, aveugles, autistes, handicaps moteurs)...
- Comment financer si le retour sur investissement est long ?
- Un Réseau RNTH ?
- Une équipe d'ingénieurs à plein temps ?

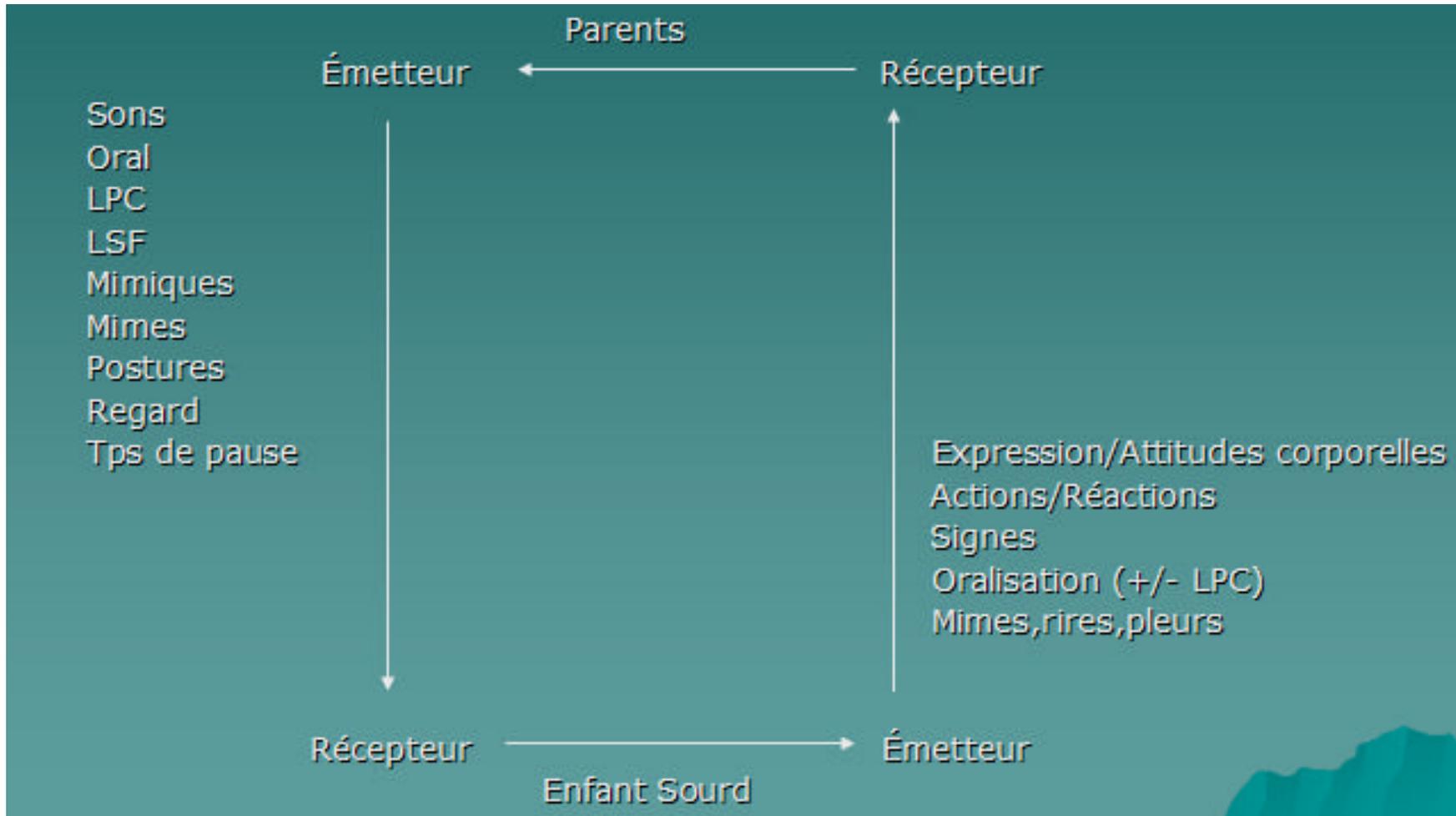
Merci de votre attention

L'équipe du projet

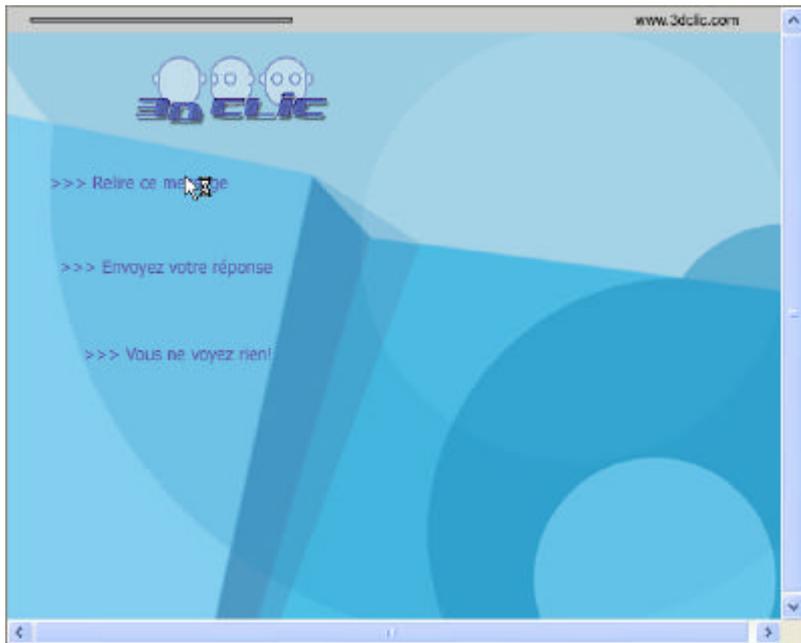


Vidéo

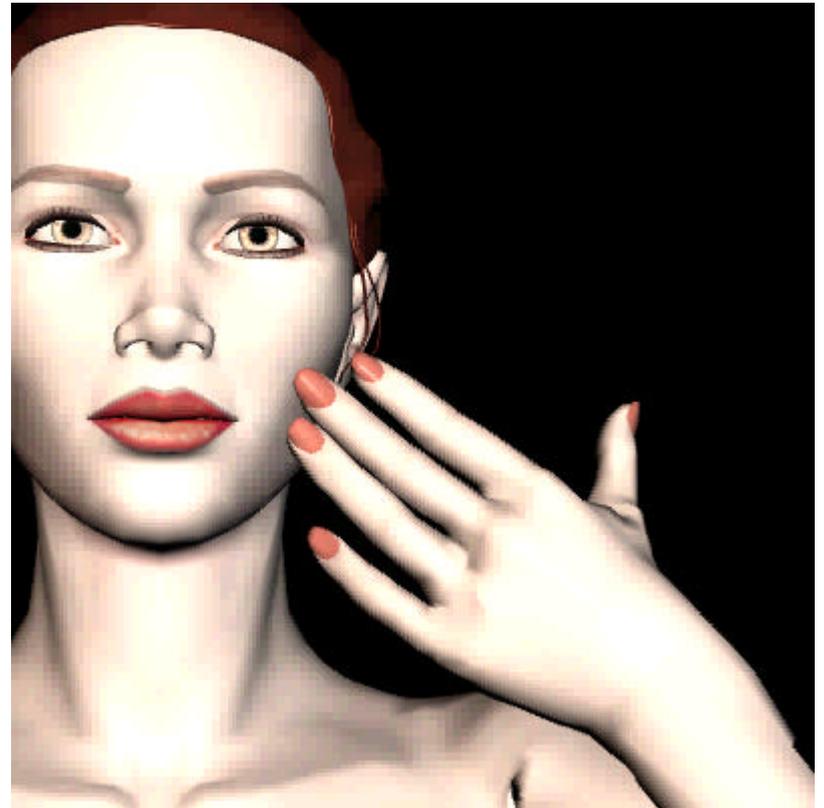
Schéma de communication



Le Message Vidéo (MMS, téléphone)

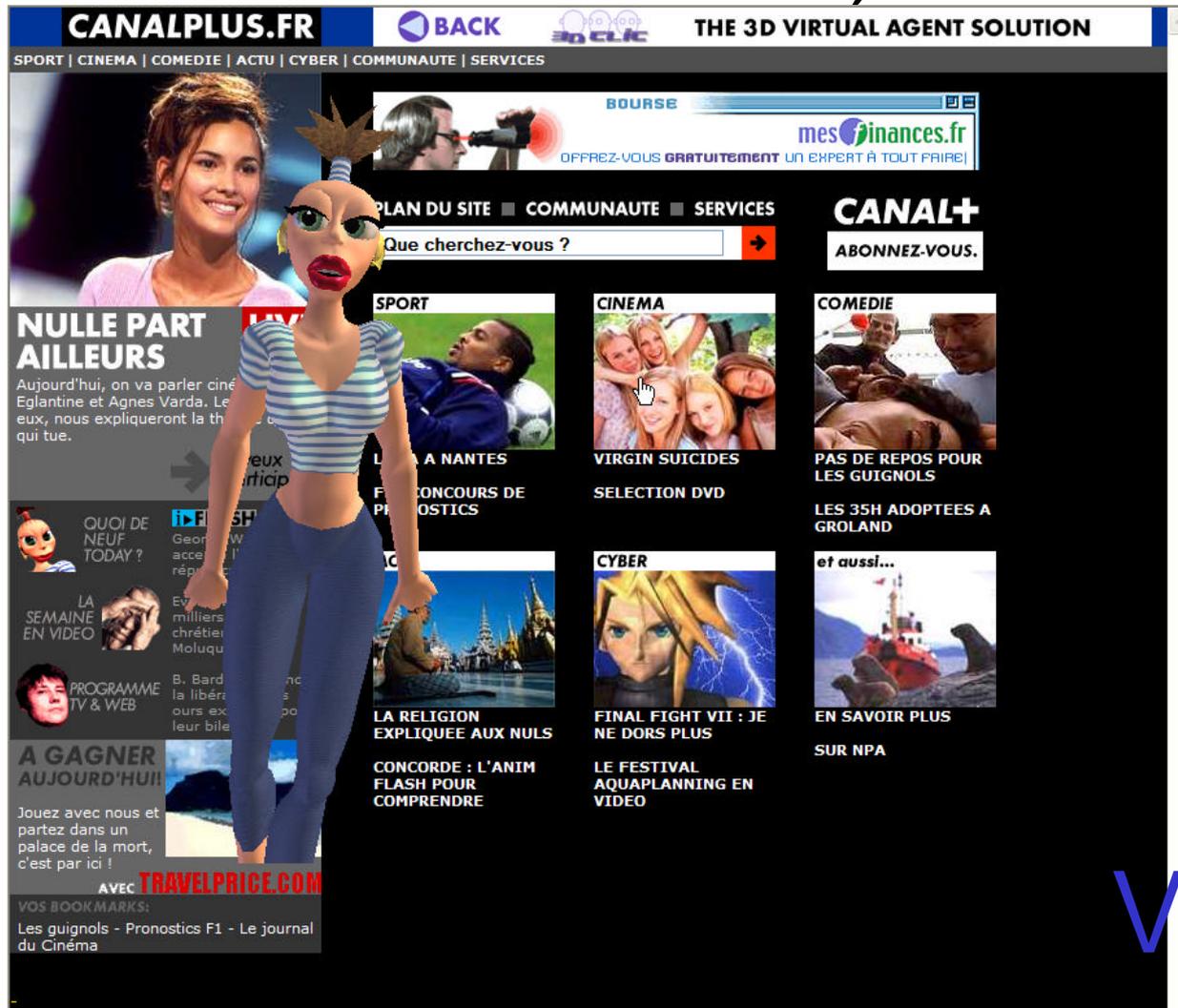


Vidéo



Vidéo

Web Marketing (navigateur, PDA, MMS...)



The image shows a screenshot of the Canal Plus website (canalplus.fr) with a 3D virtual agent overlaid on the page. The agent is a stylized female character with a large head, wearing a blue and white striped crop top and blue pants. She is positioned in the center-left of the page, partially obscuring the main content.

The website header includes the following elements:

- Navigation menu: SPORT | CINEMA | COMEDIE | ACTU | CYBER | COMMUNAUTE | SERVICES
- Logos: BACK, 3D CLIC, THE 3D VIRTUAL AGENT SOLUTION
- Search bar: "Que cherchez-vous ?" with a red arrow button.
- Buttons: "PLAN DU SITE", "COMMUNAUTE", "SERVICES", "ABONNEZ-VOUS."

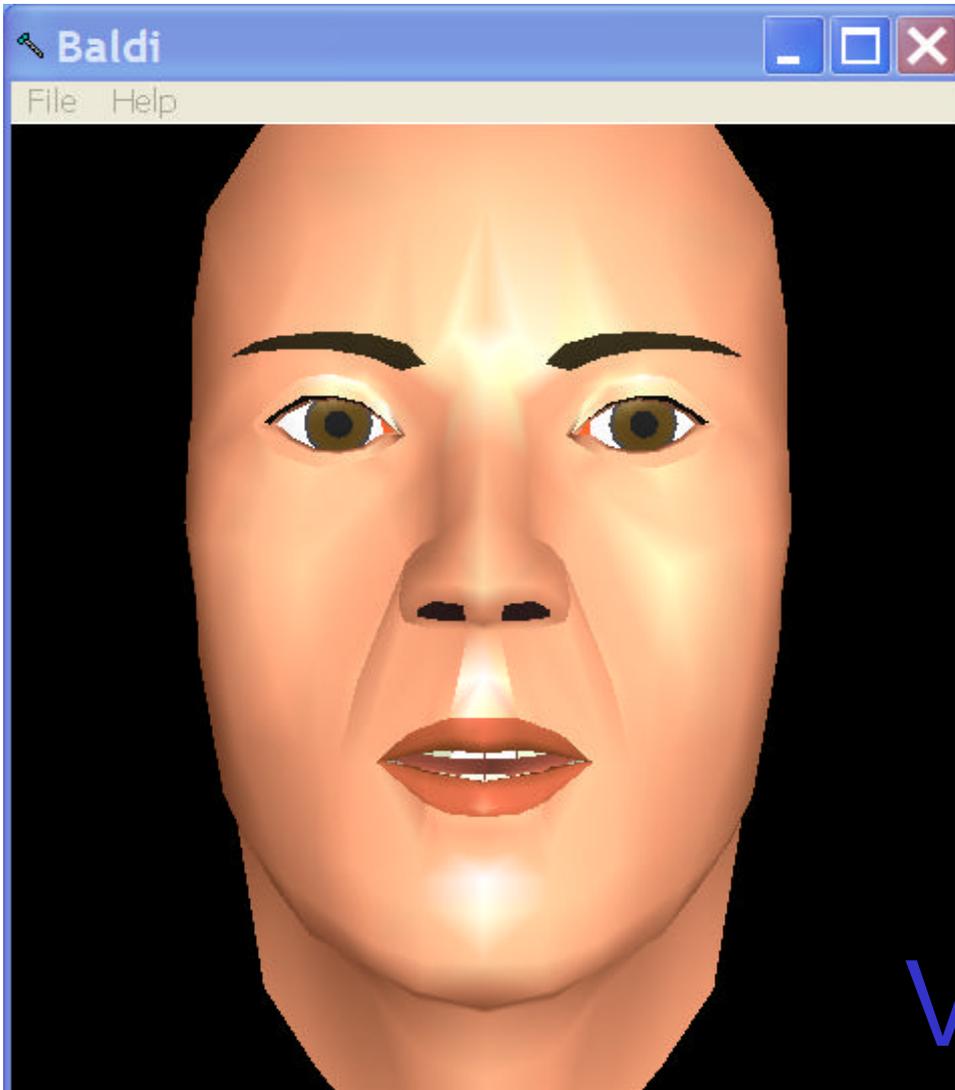
The main content area is divided into several sections:

- NULLE PART AILLEURS**: A large article featuring a woman's face and text about Eglantine and Agnes Varda.
- SPORT**: Includes "L'UNION A NANTES" and "PRONOSTICS DE F1".
- CINEMA**: Includes "VIRGIN SUICIDES" and "SELECTION DVD".
- COMEDIE**: Includes "PAS DE REPOS POUR LES GUIGNOLS" and "LES 35H ADOPTEES A GROLAND".
- CYBER**: Includes "FINAL FIGHT VII : JE NE DORS PLUS" and "LE FESTIVAL AQUAPLANNING EN VIDEO".
- et aussi...**: Includes "EN SAVOIR PLUS SUR NPA".

At the bottom of the page, there is a section for "A GAGNER AUJOURD'HUI!!" and a link to "AVEC TRAVELPRICE.COM".

Vidéo

Baldi: expressions faciales



Expressivité est importante

Reconnaissance basée sur les intonations ou des images des locuteurs

Vidéo

Version client-serveur ou multi-têtes

Téléphone, Réunion, relation clientèle

- Approfondissement de l'étude du besoin utilisateur
- Ergonomie
- Analyse et Définition architecture
 - Internet
 - Téléphone
 - Sécurité
- Développement Serveur
- Développement multi-têtes
- Développement clients
 - Navigateur web
 - Téléphone
 - PDA
- Validation
- Mise en place structure de distribution et support

Version Monoposte Texte-vers-Tête

Cds, DVds, apprentissage, web marketing

- Expression du besoin utilisateur (à approfondir)
- Graphisme tête et main (fait par Datha)
- Automate phonème visème (prototype fait par datha)
- Text to Speech (Prototype intégré par Datha)
- Moteur 3D pour remplacer Poser (à faire)
- Intégration (prototype fait par Datha)

Version Monoposte-monotête Parole-to-Tête

Classe, conférence, téléphone de bureau

- Reconnaissance du français (logiciel commercial)
- Reconnaissance de phonème (prototype INRIA ou Babel ou ?)
- Intégration (à faire)

A développer

- Un micro par locuteur
- Ou identification automatique du locuteur
- Affichage de plusieurs têtes parlantes
- Possibilité de faire ressembler la tête aux locuteurs