

Participation Francophone au Consortium C-STAR II

Hervé Blanchon, Christian Boitet, Jean Caelen

CLIPS-IMAG

BP 53

38041 Grenoble Cedex 1

herve.blanchon, christian.boitet, jean.caelen @imag.fr

1 Introduction

1.1 *La fin d'un mythe ?*

Les services d'interprétation téléphonique sont aujourd'hui en pleine expansion, et les projections économiques dans le futur montrent que, comme pour la traduction de l'écrit, on ne pourra pas former assez d'interprètes compétents. La situation classique (exemple de Hertz et AT&T aux USA) est la suivante : un client appelle un agent, et la conversation est déournée vers un interprète, qui prend alors en charge toute la conversation. Dans le futur, on imagine d'introduire entre les interlocuteurs un système de traduction automatisée, terme qui est plus modeste que le rêve sans doute irréalisable de la traduction totalement automatique (en domaine ouvert), et qui indique que les utilisateurs devront aussi aider le système (en lui parlant clairement, en le guidant, et en répondant à ses questions pour lever les ambiguïtés essentielles).

La traduction automatique de la parole a toujours été jugée comme un problème extrêmement difficile voire impossible à résoudre : d'une part reconnaître la parole continue prononcée par un locuteur quelconque, puis en comprendre le sens pour générer un énoncé dans la langue cible et enfin le synthétiser avec une voix naturelle. Autant d'étapes jugées séparément insurmontables... et pourtant, le 22 juillet 1999, une expérience de dialogue multilingue a été démontrée sur Internet, en public, devant la presse.

Cette démonstration s'est déroulée en visio-conférence avec négociations multilingues pour le renseignement touristique entre quatre sites : Grenoble, Heidelberg, Pittsburgh et Taejon. En ce qui concerne les dialogues finalisés, il s'est agi de conversations entre un agent et un client (renseignements et réservation sur un itinéraire, réservation dans un hôtel, informations et réservations touristiques).

1.2 *Les verrous scientifiques et technologiques*

Le laboratoire CLIPS-IMAG, réunit des équipes ayant des compétences en traduction automatique, reconnaissance de la parole, traitement de la langue, interfaces graphiques, qui permet à ses chercheurs de se lancer de façon crédible dans des recherches en "Traduction de parole", domaine qui est d'actualité au niveau international et qui présente de nombreuses retombées culturelles et technologiques.

Le projet tente de lever différents verrous scientifiques et technologiques :

- les verrous scientifiques : il s'agit de résoudre des problèmes difficiles de reconnaissance de la parole dans des situations peu contraintes (large vocabulaire, multilocuteurs, environnement Internet) et de traduction automatique par ordinateur (en temps réel, dans le cadre d'un dialogue). Puis dans le sens contraire, de génération dans la langue cible et de synthèse vocale.

- les verrous technologiques : le système devant être utile à des usagers potentiels; il est nécessaire de bâtir un plan expérimental pour étudier le comportement de sujets devant communiquer en diverses langues au moyen d'un ordinateur traducteur. Pour cela il a fallu monter un laboratoire numérique de « capture » et d' « observation » en technique de Magicien d'Oz (en simulant les fonctionnalités de la machine à l'aide de compères) pour réunir les matériaux de l'étude : corpus, annotations, etc. Puis, il a fallu réaliser un système de communication sous Internet et enfin il a fallu valider et évaluer l'utilisabilité du système final avec des utilisateurs en vraie grandeur.

1.3 Problèmes scientifiques abordés

Les principaux problèmes de reconnaissance ou de traduction demeurent dès que l'on s'éloigne de situations calibrées et contraintes. C'est particulièrement le cas de la parole spontanée :

- elle présente des inattendus et des irrégularités linguistiques,
- elle présente des reprises, des hésitations, des bruits parasites, etc.,
- la variabilité des signaux est croissante avec les conditions acoustiques de prise de son ou d'environnement (effet de bruit, effet du canal de transmission, etc.),
- le nombre important de locuteurs pose des problèmes d'identification, d'adaptation, etc.

Les problèmes fondamentaux en traduction demeurent aussi lorsque le domaine du discours s'élargit et que l'on prend le cadre du dialogue en plus, comme c'est le cas ici. Les problèmes qui s'ajoutent à la traduction proprement dite sont des problèmes d'analyse et de compréhension de la langue ainsi que des problèmes de conduite du dialogue. Parmi ceux on notera les problèmes de référent, les ambiguïtés, les anaphores et la résolution des déictiques dans le dialogue.

Enfin une troisième classe de problèmes apparaît pour le choix des stratégies d'aide du dialogue. Le système doit non seulement traduire le « sens » de l'énoncé mais aussi doit redonner « l'intention » sous-jacente. Entrent ici en ligne de compte les forces illocutoires et perlocutoires c'est-à-dire la manière dont un énoncé doit être dit pour conseiller, suggérer, ordonner, énoncer, etc. C'est cette dernière classe de problèmes qui rend difficile la génération et la synthèse de la parole en situation de dialogue : en effet il ne s'agit plus de lire un texte généré, mais bien de traduire avec la même intonation ou du moins la même *intention* d'intonation, les énoncés du locuteur.

2 Le consortium C-STAR

2.1 C-STAR I (91-93)

ATR (Advanced Telecommunication Research) a été créé en avril 1986. Il s'agit d'une structure de droit privé, financée à 70% par le ministère des finances du Japon et à 30% par un consortium de 200 sociétés, au premier rang desquelles NTT et KDD. Les buts d'ATR ne sont pas seulement scientifiques. Il s'agit d'une opération de prestige, engagée dans le cadre de l'internationalisation, thème moteur officiel du Japon des années 80. C'est pourquoi ATR-ITL (Interpreting Telecommunications) s'est voulu le moteur d'un consortium international visant en particulier à réaliser des démonstrations spectaculaires et bien médiatisées à la fin de chaque projet.

CSTAR-I est le nom du consortium qui a monté les premières démonstrations publiques d'interprétation téléphonique en janvier 1993 dans le cadre du premier projet ATR-ITL. Les partenaires en étaient ATR, CMU (Pittsburgh, USA), Siemens (SNI, Munich), et l'université de Karlsruhe.

Devant le succès remporté par la première démonstration, le projet de ATR-ITL a été renouvelé et les fondateurs de C-STAR-I ont décidé de poursuivre leur collaboration en essayant d'impliquer d'autres groupes de recherche. Dès 1993, ETRI (Electronics and Telecommunications Research Institute), un puissant institut coréen créé un peu à l'image de ATR a rejoint le nouveau consortium.

2.2 C-STAR-II (93-99)

Le consortium C-STAR II est une organisation souple qui réunit des groupes de recherche qui veulent bien travailler ensemble autour d'un thème commun : la traduction de parole. Le consortium n'est pas une source de financement. Les participants doivent trouver leurs propres sources de financement. Deux types de participations sont possibles, en tant que partenaire ou en tant qu'affilié. Les partenaires s'engagent à construire un système complet qui leur permet de participer aux démonstrations finales en traduisant vers au moins une langue. Les affiliés font de la recherche dans le domaine et participent à ses avancées, mais ne s'engagent pas à construire un système complet. Au niveau des partenaires, l'IRST (Istituto per la Ricerca Scientifica e Tecnologica) à Trento en Italie et le groupe CLIPS++¹ sont venu enrichir le nombre de langues disponibles. Le schéma suivant donne la liste et la distribution géographique des partenaires (en gras) et des affiliés.

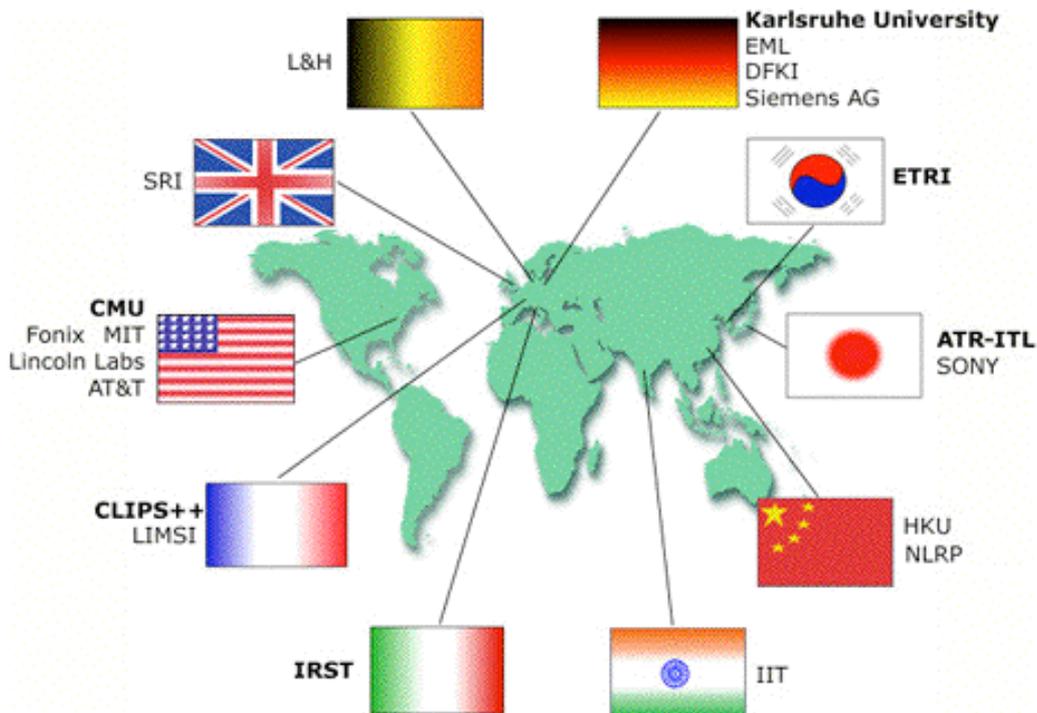


Figure 1 : Les partenaires et les affiliés de C-STAR II

2.3 Approches possibles

Un prototype C-STAR II doit fournir les services suivants pour mener à bien un dialogue entre deux partenaires ne parlant pas la même langue :

- En entrée
 - Un module de reconnaissance de la parole qui produit un texte plus ou moins annoté dans la langue de celui qui parle
 - Un module qui traduit le texte de celui qui parle
 - Soit vers une structure pivot que l'on appelle IF. On parle d'un analyseur vers l'IF.
 - Soit vers une (ou plusieurs) langue(s) cible(s). On parle d'un module de traduction.
- En sortie
 - Un module de génération qui traduit l'IF de celui qui vient de parler en un texte dans la langue de celui qui écoute. Si un site reçoit un texte dans sa propre langue, il n'a rien à faire.
 - Un module de synthèse vocale pour prononcer le texte dans la langue de celui qui écoute.

¹ Cf. Infra pour la constitution du groupe

Le schéma ci-dessous illustre les deux chemins possibles.

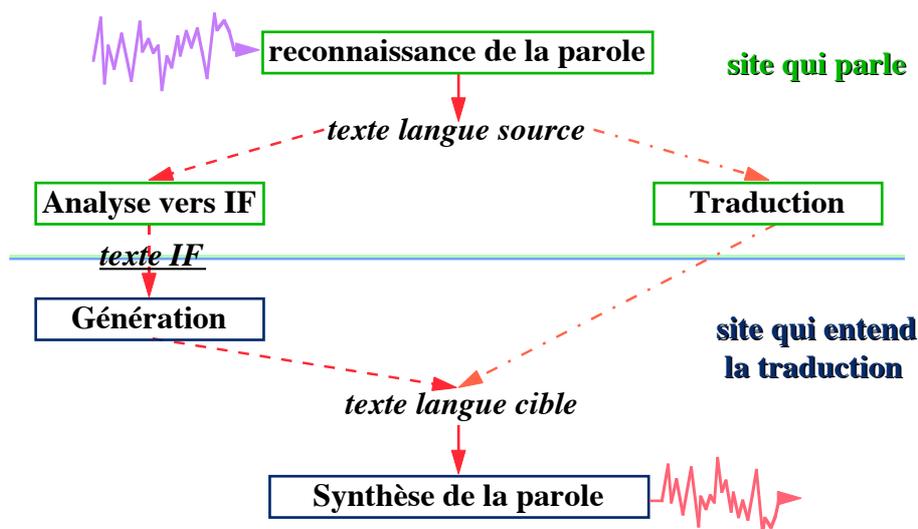


Figure 2 : Traduction par pivot vs traduction directe

L'IF (Interchange format) est une structure sémantique indépendante des langues. Elle est fondée sur :

- des actes de dialogue : ce que l'on veut ou ce que l'on fait give-information
- des concepts : à propos de quoi (focus informationnel) +availability+room
- des arguments : valeurs des variables du discours room-type=(single ; double)

Par exemple, pour la phrase "la semaine du 12 nous avons des chambres simples et doubles disponibles" prononcée par un agent de voyage francophone, l'IF produite est :

a:give-information+availability+room(room-type=(single ; double), time=(week, md12))

La correspondance entre les mots de la phrase et l'IF s'établit comme suit :

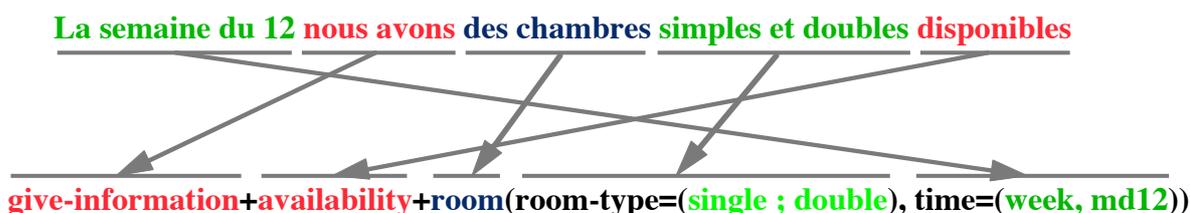


Figure 3 : Correspondance entre un énoncé en français et l'IF associé

L'avantage de cette technique pour celui qui l'utilise est qu'elle permet avoir comme langue cible toutes les langues passant par l'IF. Cette technique est utilisée par tous les partenaires du consortium sauf le Japon. De plus, un partenaire qui passe par l'IF à aussi un générateur à partir de l'IF pour écouter les autres partenaires. Celui qui parle peut donc utiliser son propre générateur pour régénérer (sous forme textuelle seulement) l'IF que le système a produite pour ce qu'il vient de dire. Ainsi, l'utilisateur peut voir ce que le système a compris pour peu que le générateur soit "fiable". Par "fiable" on entend que pour une IF donnée, le générateur produit une paraphrase exprimant exactement le contenu de l'IF. Le schéma suivant illustre la situation pour le français.

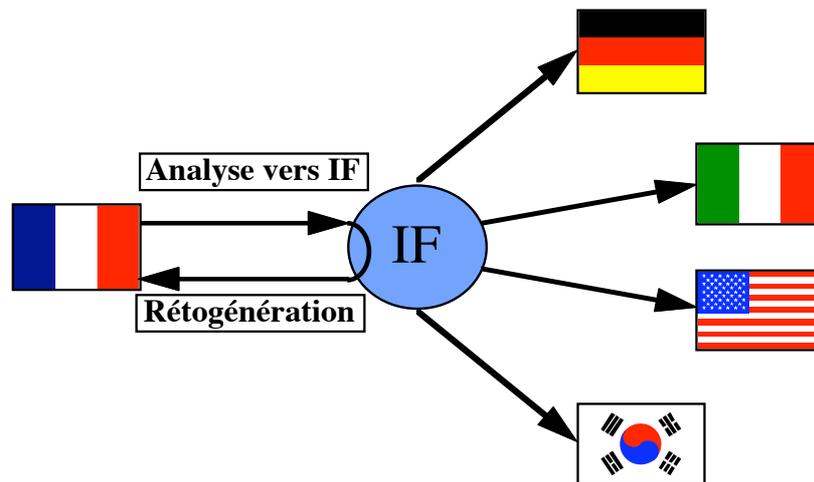


Figure 4 : Analyse vers IF pour aller vers les langues cibles et rétro-génération pour contrôler le sens émis vers les récepteurs

3 Contribution du partenaire CLIPS++

3.1 L'équipe francophone

Le partenaire francophone CLIPS++, coordonné par le laboratoire CLIPS (Communication Langagière et Interaction Personne Système - IMAG, France), se compose du LATL (Laboratoire d'Analyse et de Technologie du Langage - Université de Genève, Suisse), du LAIP (Laboratoire d'Analyse Informatique de la Parole - Université de Lausanne, Suisse), et du LIRMM (Laboratoire d'Informatique, Robotique et Microélectronique de Montpellier - Université de Montpellier, France) a rejoint le consortium C-STAR II en tant que partenaire en Septembre 1996.

Le laboratoire CLIPS, réunit en effet des équipes ayant des compétences en traduction automatique, reconnaissance de la parole, traitement de la langue, interfaces graphiques, qui nous a permis de nous lancer de façon crédible dans des recherches en traduction de parole.

Le CLIPS++ a finalement choisi de passer par l'IF pour la traduction ce qui nous permet de bénéficier des avantages exposés ci-dessus. Une expérience est cependant en cours avec le LATL pour comparer le passage par l'IF et la traduction directe vers l'anglais.

3.2 Les composants du groupe CLIPS++

Les composants suivants ont été développés :

- reconnaissance de la parole par le groupe CLIPS-GEOD
- français vers IF par le groupe CLIPS-GETA
- IF vers français par le LATL
- synthèse du français par le LAIP
- intégrateur et interface par le groupe CLIPS-MULTICOM

Voici quelques informations sur ces différents composants.

3.2.1 Reconnaissance de la parole

C'est un module de reconnaissance multilocuteur de parole continue spontanée avec un vocabulaire centré sur le domaine du tourisme de 10k mots. Il est basé sur une architecture client-serveur, c'est à dire que le serveur de reconnaissance est mis en œuvre via des clients "légers" sur le réseau. Ce

module est construit sur la boîte à outils de JANUS III en collaboration avec l'Université Carnegie Mellon.

Sont mis en œuvre pour produire le texte correspondant à une entrée vocale :

- Un modèle acoustique markovien indépendant du contexte entraîné sur 10 heures de parole continue (corpus BREF-80).
- Un modèle de langage stochastique entraîné sur un corpus de 140 millions de mots et optimisé pour la tâche de renseignement et de réservation touristique.

3.2.2 Français vers IF

Ce module est développé sous Ariane-G5 qui est un générateur de systèmes de traduction qui utilise cinq langages spécialisés pour la programmation linguistique dans l'environnement VM/ESA/CMS. Il est actuellement opérationnel sur trois sites : à Grenoble (sur un IBM 9221-130 à 3,5 mips), au CCSJ Marseille (sur un IBM 9672-R14 à 40 mips) et à Montpellier au centre de calcul IBM (sur un IBM 9672-RX5 à 60 mips).

En entrée ce module reçoit une transcription orthographique de l'énoncé oral. Les étapes suivantes se succèdent jusqu'à obtenir en sortie, un texte en langue IF :

- Analyse morphologique et lemmatisation des mots du texte,
- Première consultations de dictionnaires de transfert vers l'IF
- Analyse syntaxique pour la reconnaissance de structures grammaticales sémantiquement pertinentes : dates, quantités, numéros, prix
- Autres consultations de dictionnaires de transfert vers l'IF
- Génération syntaxique puis morphologique de l'IF

3.2.3 IF vers français

Ce module est développé en partie sous GBGen qui est un outil de génération syntaxique à large couverture lexicale et grammaticale. C'est un outil déterministe basé sur une grammaire générative

Le passage d'une IF à une forme orthographique du français se fait en trois étapes :

- mise en correspondance entre une structure IF et une structure sémantique de GBGen
- mise en œuvre des procédures de génération de GBGen pour produire une structure syntaxique
- mise en œuvre de règles morphologiques dans GBGen pour produire une forme orthographique

3.2.4 Synthèse du français

Le module LAIPTTS est un synthétiseur "text-to-speech" basé sur des règles. La synthèse se déroule en trois étapes : – conversion "texte vers phonèmes", – génération de la prosodie, – génération du signal.

La conversion "texte vers phonèmes" se fait avec des règles générales (540) et spécialisées pour les nombres, les abréviations, les expressions figées, etc. Elle met en œuvre des dictionnaires général (7000 mots) et spécialisé (noms propres). La génération de la prosodie met en œuvre des règles de type psycho-linguistique. La génération du signal utilise la technique Mbrola de la faculté de Mons en Belgique.

3.3 Le démonstrateur du CLIPS++

3.3.1 Communication entre les démonstrateurs

Les échanges entre les démonstrateurs sont de deux types : les échanges audio et vidéo assurés par des systèmes de visio-conférence du commerce, et les échanges concernant le processus de traduction proprement dit.

Les échanges audio et vidéo se font en connexion point à point si deux sites seulement souhaitent communiquer et via un pont si plusieurs sites souhaitent communiquer.

Les échanges concernant le processus de traduction se réduisent au minimum et consistent en des chaînes de caractères. Ces échanges se font au moyen d'un serveur de communication sur lequel se connectent les sites qui veulent communiquer. La connexion au serveur de communication se fait via le protocole Telnet. Cela permet à tous les sites connectés sur le serveur de communication de recevoir (d'entendre) de qui est émis (dit) par l'un des sites.

Le serveur de communication inter-sites est appelé serveur de communication global.

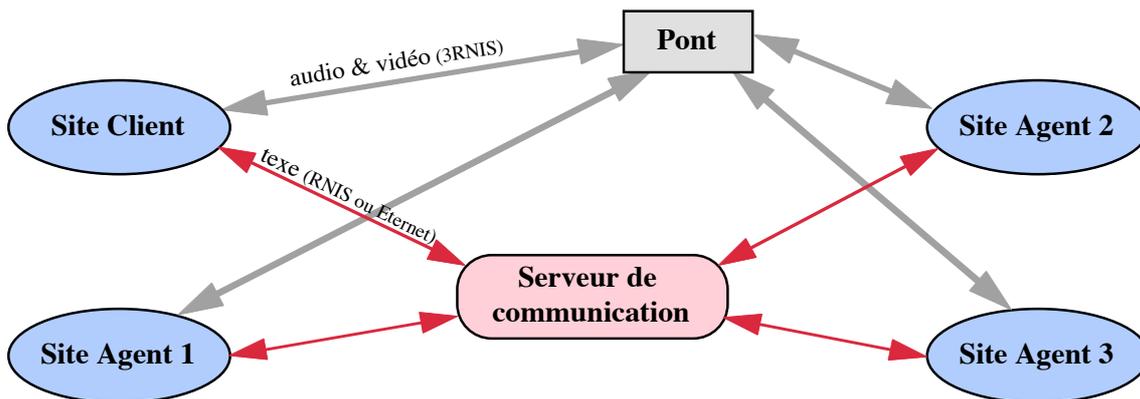


Figure 4 : Les échanges de visio-conférence et de traduction

3.3.2 Communication entre les composants du démonstrateur CLIPS++

Dans le démonstrateur du CLIPS++ les composants agissent tous comme des serveurs. Ainsi, aucun ne communique jamais directement avec un autre. Nous avons utilisé une réplique de l'architecture globale de communication entre sites distants pour la communication entre les composants locaux en mettant en place un serveur de communication local.

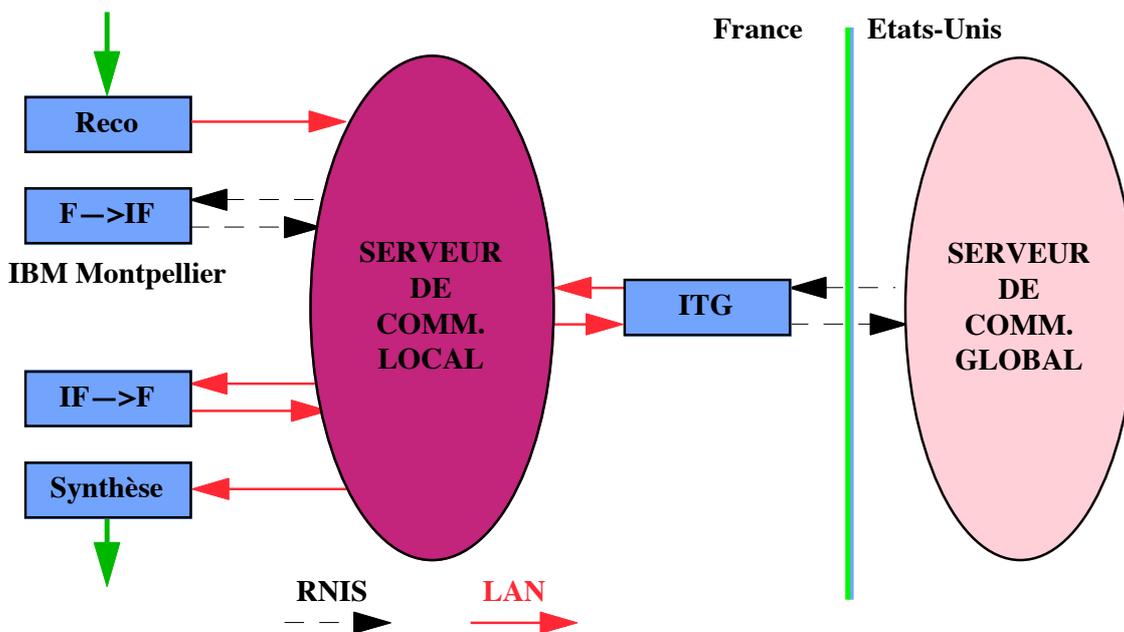


Figure 5 : Architecture local du démonstrateur du CLIPS++

Chacun des composants (serveurs) reçoit tous les messages qui sont émis sur le serveur de communication local. Les messages qui transitent sur le serveur de communication local se composent de quatre informations :

- le type de la donnée (hypothèse de reconnaissance, IF, texte de génération),
- la langue de la donnée (français, allemand, anglais, coréen),
- le nom de l'émetteur du message et un identificateur unique de message,
- la donnée que véhicule le message (texte de reconnaissance, texte IF, texte de génération)

Les composants savent filtrer les messages et ne répondent qu'à ceux qui le concernent en fonction du type de données reçu. Une fois le traitement effectué et si nécessaire, chacun des composants, envoie un message vers le serveur de communication local qui contient le type de données produit et le résultat du traitement.

Cette architecture permet un développement séparé des modules, le remplacement d'un module par un autre, et la distribution des modules qui peuvent résider physiquement sur des machines, des systèmes d'exploitations et des sites différents. Le coût des communications étant très peu coûteux en temps, il ne nuit pas à la performance globale du démonstrateur.

L'intégrateur permet de transmettre seulement les données pertinentes sur le serveur de communication global, en particulier les IFs qui correspondent à ce qui est dit en français, et de récupérer les IFs émises par les sites distants qui correspondent à ce qui a été dit sur ces sites.

Deux types de liaisons réseau sont mise en œuvre pour faire coopérer et échanger les modules. Des liaisons Ethernet et des liaisons RNIS. Les liaisons RNIS servent seulement à fiabiliser les échanges sur de longues distances. L'expérience a montré que ces dernières sont inutiles et que, étant donnée la faible taille des données échangées, des liaisons Ethernet auraient suffi.

3.3.3 Interface du démonstrateur

L'interface sur démonstrateur est distribuée sur deux écrans : le premier sert à afficher la visioconférence entre le client et les différents agents de voyage, le second est dédié à l'interface utilisateur proprement dite.

L'écran d'interface (cf. annexe 1) est divisé en deux. À gauche se trouve l'interface du système de traduction proprement dit. La partie droite est consacrée à l'affichage des pages web que l'agent de voyage propose au client.

L'interface du système de traduction est détaillée dans l'annexe 2. Le bouton PTT est utilisé pour contrôler le module de reconnaissance de la parole. Lorsqu'il est enfoncé, le bouton devient orange, indiquant que le module de reconnaissance est prêt. Lorsque l'utilisateur commence à parler, le bouton devient vert, la reconnaissance de ce qui est dit commence et les résultats intermédiaires sont proposés à l'utilisateur en bleu. Lorsque l'utilisateur se tait, le bouton devient de nouveau gris et le résultat final de la reconnaissance est affiché en noir. Nous aurions ainsi pu nous passer de la technique du Push To Talk. Mais nous avons choisi de ne pas produire d'IF pour des reconnaissances trop éloignée du tour de parole d'origine. C'est une solution qui permet de gagner du temps puisque dans le démonstrateur actuel le module français -> vers IF est moins rapide que le module de reconnaissance de la parole. Nous utilisons donc un bouton de validation d'hypothèse de reconnaissance.

Lorsque l'utilisateur est un agent, l'interface comprend aussi un module qui permet de choisir une page web à envoyer au client. Les pages envoyées sont aussi affichées localement pour le contrôle visuel de l'agent.

4 La démonstration du 22 juillet 1999

4.1 Lieu, audience et couverture médiatique

La démonstration s'est déroulée dans les locaux de la maison Jean Kuntzmann à l'institut IMAG devant 90 personnes parmi lesquelles des représentants des tutelles du CLIPS, des représentants de l'industrie et de centres de recherche (IBM, Xerox, Digigram, Azimut, FranceTelecom, Magellan Ingénierie, CNET, Spacio Guide) ainsi que de l'ANVAR.

Nous avons aussi fait des démonstrations privées de tout ou partie du système à Alcatel, Thomson, Texas Instrument, Schneider, Parrot, CyberStudio, EDF, Soprane, Grany Media.

La couverture médiatique fut assez bonne le 22 juillet et la semaine suivante. Il existe encore un certain intérêt pour nos résultats et de nouvelles publications doivent avoir lieu.

La situation est pour l'instant la suivante :

- Agences de Presse : AFP, AGRAP, Temps d'm, Agence REA, Editing Corporate.
- Presse écrite : Le Monde (Interactif), Libération, Le Dauphiné Libéré, Science & Vie, Science & vie junior, Micro-Hebdo, et à paraître, Science & Vie Micro (dossier) et Wired.
- Télévision : France 3. Pour être un peu polémique, nous pouvons constater que la télévision n'a pas été au rendez-vous bien que conviée comme les autres médias.
- Radio : Radio France Isère, France Inter, France Info, BFM, Europe 2, RMC, RTL, RCF Isère.

4.2 Scénario démontré

La fin de matinée du 22 a été consacrée à une session introductive à propos de C-STAR II, son histoire, les techniques mises en œuvre dans la traduction de parole, et le démonstrateur du groupe CLIPS++. La démonstration en elle même a eu lieu à 14 heures.

Nous avons montré le scénario suivant :

- Un client arrive dans une agence de voyage "virtuelle" avec des relais locaux en Corée du Sud, en Allemagne et aux États-Unis. Il salue d'abord les trois agents de voyage disponibles et leur demande des informations sur l'heure, le temps et la température dans leurs villes respectives. Dans cette phase toutes les synthèses des agents de voyage étaient renvoyées dans le circuit de visio-conférence. Ainsi le client entendait en même temps les traductions en coréen, anglais et allemand de ce qu'il venait de dire. Les étapes suivantes se déroulaient avec la seule synthèse de l'agent interlocuteur audible.
- Le client organise ensuite un voyage avec un collègue à Taejon en Corée du Sud. Il achète des billets d'avion pour Séoul, s'enquiert des moyens de transport possibles de Séoul à Taejon, réserve une chambre d'hôtel, demande des informations touristiques sur les environs de Taejon, et règle avec une carte de crédit.
- Le client organise alors un voyage à New-York avec sa femme. Il achète des billets d'avion pour New-York, réserve une chambre d'hôtel, demande à voir un match de base-ball, et règle avec une carte de crédit.
- Le client organise finalement un voyage à Heidelberg. Il achète un billet de train, cherche une chambre d'hôtel à un prix raisonnable, demande des informations touristiques sur les environs de Heidelberg, et règle avec une carte de crédit.
- Pour les besoins de la démonstration, le client salue enfin nominativement les trois agents avec lesquels il a interagi.

Nous avons aussi participé aux démonstrations de ETRI en faisant l'agent de voyage français.

4.3 Résultats

Bien que nous ayons réalisé tardivement que le scénario général était trop répétitif, l'audience n'a fait aucun commentaire sur ce sujet. Nous avons cependant bien veillé à utiliser des formulations différentes chaque fois les mêmes thèmes étaient évoqués. La cérémonie d'ouverture, avec les synthèses distantes fonctionnant en parallèle à été très bien reçue. Quelques erreurs ont su amuser les spectateurs. La démonstration a duré une demie heure pendant laquelle les gens n'ont pas vu le temps passé.

La conversation avec l'agent coréen a eu le plus gros succès (l'écriture hangul et une langue que l'on n'entend jamais) et a bien montré l'intérêt de la traduction de parole lorsqu'il y a un message à échanger et pas de langue en partage pour le faire. En Europe, les gens sont moins sensibles à cet état de fait en ce qui concerne l'allemand ou l'anglais. La plupart des exemples repris par les médias concernaient des applications potentielles de la traduction de parole français japonais ou français coréen. Nous avons pourtant essayé le faire passer le message que même si une langue est en partage, pour des choses qui dépassent le vocabulaire de tout les jours, une assistance est de toute façon nécessaire pour faire passer un message qui doit être précis.

La vitesse des systèmes n'a pas été commentée. Tout a semblé se passer à une vitesse raisonnable et acceptable. Dans le démonstrateur du CLIPS++ il faut en moyenne 5 secondes pour faire du français vers l'IF. L'IF vers le français est plus rapide, de l'ordre de 1 à 2 secondes. L'échange de pages web contenant parfois des visites virtuelles des hôtels ou des musées a été aussi bien apprécié. Cela semble faire maintenant parti du décor et enrichir grandement l'interaction.

5 Conclusion et perspectives

Cette expérience nous a permis de vérifier que nous pouvions atteindre l'objectif des démonstrations communes en un temps record. Et elle a mis en évidence un grand nombre de problèmes scientifiques intéressants à résoudre. Il faut s'intéresser au suivi du dialogue et la prédiction qui sont indispensables pour résoudre les problèmes d'élosion et de références qui sont partout présents dans le dialogue. La représentation IF est intéressante, mais pour mieux faire, il faut qu'elle permettent de représenter plus d'informations qui rendront les échanges plus riches et transmis avec plus de précisions.

6 Références

- [1] Boitet, Ch. (1998) Problèmes scientifiques intéressants en traduction de parole. Proceedings of TAL+AI/NLP+IA, Moncton, Canada. Septembre 98.
- [2] Boitet, Ch. (1997) GETA's MT methodology and its current development towards networking communication and speech translation in the context of the UNL and CSTAR projects. Proceedings of PACLING-97, Ome, Japan. 2-5 September 97.
- [3] Keller, E., & Zellner, B (1998). Motivations for the prosodic predictive chain. Proceedings of ESCA Symposium on Speech Synthesis. Paper 76, pp. 137-141. Jenolan Caves, Australia.
- [4] Keller, E. (1997). Simplification of tts architecture vs. Operational quality. Proceedings of EUROSPEECH '97. Paper 735. Rhodes, Greece. September 1997.
- [5] Keller, E., & Werner, S. (1997). Automatic Intonation Extraction and Generation for French. 14th CALICO Annual Symposium. ISBN 1-890127-01-9, West Point. NY. June 1997.
- [6] Vaufray D., Akbar M., Rouillard J. (1999) A Network Architecture for Building Application that Use Speech Recognition and/or Synthesis. Proceedings of Eurospeech'99, Budapest, pp 2159-2162
- [7] Wehrli, E. & Wehrle T (1998) Overview of GBGen. Proceedings of 9th International Workshop on Natural Language Generation, Niagara-on-the-lake, Canada. August 1998.
- [8] Wehrli, E., Wehrle T., Mengon J., Vandeventer A. (1999) Une approche efficace à la génération syntaxique. Le système GBGen. Proceedings GAT'99, Grenoble, France. octobre 1999