Traduction de la parole pour le français : une première étape et quelques perspectives

Speech Translation for French: a First Demonstrator and Some Prospects

Hervé Blanchon, Christian Boitet & Jean-Philippe Guilbaud

GETA-CLIPS
BP 53
38041 Grenoble Cedex 9
{herve.blanchon, christian.boitet, jean-philippe.guilbaud}@imag.fr

Résumé

Partenaire du consortium C-STAR II, le groupe CLIPS++ a participé à des démonstrations intercontinentales de traduction de dialogues oraux finalisés dans le domaine du tourisme le 22 juillet 1999. Ces démonstrations ont aussi impliqué CMU (Etats-Unis), ETRI (Corée du Sud), ATR (Japon) et UKA (Allemagne). Pour ce faire, nous avons dû construire les modules de traitement du français dans un temps relativement bref à partir de la fin de l'année 1996. Cet article se propose de donner une vue d'ensemble du travail du groupe CLIPS++.

Nous présentons d'abord l'architecture et les différents modules de la partie française du démonstrateur, puis nous évaluons nos résultats tant du point de vue interne qu'externe. Bien que les réactions aux démonstrations aient été très positives et que l'opinion généralement émise était que la mise sur le marché devait pouvoir se faire à très court terme, nous pensons qu'il reste bien du chemin à parcourir pour améliorer la qualité de manière significative.

Dans la dernière partie, nous proposons des pistes de recherche qui doivent permettre d'améliorer la traduction de dialogues oraux finalisés en particulier en définissant un pivot sémantique mieux construit et plus puissant, en utilisant le contexte linguistique et dialogique, et en produisant des informations utilisables par un synthétiseur de parole afin d'obtenir une vraie prosodie du dialogue.

Mots Clés

Traduction de parole, C-STAR, Pivot sémantique

Abstract

As a partner of the C-STAR II consortium, the CLIPS++ group took part in multilingual, intercontinental demonstrations of speech translation within the tourism domain held on July 22nd 1999 by CLIPS (France), CMU (United States), ETRI (South Korea), ATR (Japan), IRST (Italy), and UKA (Germany). Starting at the end of 1996,we had to build the modules to handle French within a short time. In this paper we propose an overview of our experience.

After presenting the modules and the architecture of the CLIPS++ demonstrator, we evaluate the results, both externally and internally. While the reactions to the demonstrations were very positive, and many said that these prototypes should quickly lead to products, we feel that there is still much room for improving the overall quality in significant ways.

In the last part, we focus on future avenues of research to further improve the quality of task-oriented speech translation, in particular by defining a more powerful and orthogonal task-oriented semantic pivot, using the linguistic and dialogic context, and generating information usable by speech synthesis to reach a dialogue oriented prosody.

Keywords

Speech Translation, C-STAR, task-oriented semantic pivot

1. Introduction

Le groupe CLIPS++ a rejoint le consortium international C-STAR II[1] (Consortium For Speech Translation Advanced Research), en tant que partenaire, en septembre 1996. Coordonné par le laboratoire CLIPS (France), notre groupe était composé de trois autres laboratoires LATL (Suisse), LAIP (Suisse) et LIRMM (France). En parallèle aux activités du consortium, le projet VerbMobil[2] en Allemagne s'intéresse aussi à la traduction de parole.

Nous avons adopté l'approche interlingue, déjà utilisée par 4 des partenaires C-STAR II, dans laquelle l'interlingue, appelé IF, est un pivot sémantique spécialisé pour la tâche. Nous avons donc réalisé 4 modules : reconnaissance de la parole pour le français, analyse du français vers l'IF, génération du français à partir de l'IF, et synthèse orale du français. Ces modules coopèrent entre eux et avec les démonstrateurs des autres partenaires via un intégrateur. Le scénario envisagé dans le domaine du tourisme permet à un client d'acheter des titres de transport, de réserver un hôtel, et d'avoir des informations touristiques sur le lieu de destination.

La traduction de parole de faible qualité, mais néanmoins utile, peut être réalisée en mettant bout à bout des systèmes de reconnaissance, de TA, et de synthèse du commerce. C'est l'approche "quick and dirty" [Seligman & al., 98] dans le contexte de conversations informelles et épisodiques, telles que celles des forums de discussion, démontrée par Marc Seligman à différentes reprises. Il existe cependant des situations pour lesquelles la qualité est indispensable, celles qui ont un caractère d'urgence ou celles qui implique au moins un « agent » professionnel à la disposition d'un client. Ce dernier cas est celui où nous nous plaçons en mettant en contact des clients avec des agents de voyage.

De fait, l'approche par pivot sémantique spécialisé pour la tâche et les techniques heuristiques de programmation linguistique utilisées dans le cadre du projet permettent d'ajouter de nouvelles langues facilement et d'obtenir une qualité de traduction qui est manifestement meilleure.

Avec l'approche "quick and dirty" il faut parfois répéter 4 ou 5 fois un tour de parole avant qu'il soit reconnu comme il faut, ou édité manuellement de telle façon que le processus complet de traduction prend de 20 à 30 secondes. Lors des démonstrations C-STAR, on observe rarement plus d'une répétition et la traduction prend 4 à 7 secondes, sans que l'utilisateur ait l'impression d'attendre du fait de l'environnement de visioconférence multimédia utilisé.

De plus, nous pensons qu'il reste pas mal de chemin à parcourir pour améliorer encore la qualité des systèmes de traduction de dialogues oraux finalisés.

2. Composants développés

L'IF [Levin & al. 98] est basé sur des actes de dialogues, des concepts et des arguments. Les actes de dialogue décrivent les intentions, les besoins de celui qui parle `(give-information, introduce-self…)`. Les concepts définissent à propos de quoi les actes de dialogue sont exprimés `(availability, train, flight, room…)`. Les arguments permettent d'instancier les valeurs des variables du discours `(room-type, time…)`. L'IF suivante : `a:give-information +availability+room (room-type=(single ; double), time= (week, md12))` correspond à un énoncé produit par un `agent` dont le

[1] Voir aussi [Lavie & al. 2000, Park & al. 98, Sugaya & al. 99, Sumita & al. 99, Takezawa & al. 98]

[2] Consulter http://verbmobil.dfki.de/ pour des informations complémentaires.

contenu sémantique est « la semaine du 12, nous avons des chambres simples et double disponible ».

L'architecture d'un système de traduction utilisant l'approche pivot est décrite Figure 1 page suivante.

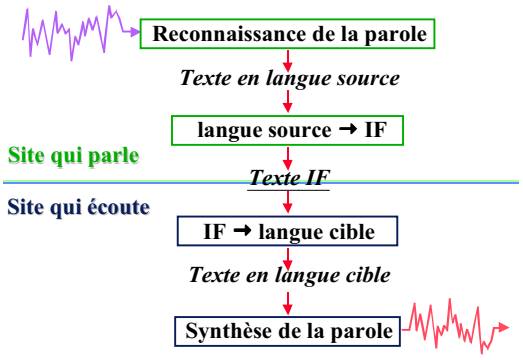


Figure 1: Traduction via l'IF

Les modules reconnaissance de la parole et français vers IF ont été développés au sein du CLIPS, respectivement par les équipes GEOD et GETA. Le module IF vers français à été développé au sein du LATL. Le module de synthèse a été développé au sein du LAIP. Le LIRMM a pour sa part exploré la réalisation d'un module français vers IF avec une autre méthodologie.

2.1. La reconnaissance de la parole

C'est un module [Vaufreydaz & al. 99] de reconnaissance multilocuteurs 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 a 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.

Pour produire une transcription orthographique correspondant à une entrée vocale, il met en œuvre un modèle acoustique markovien indépendant du contexte entraîné sur 10 heures de parole continue (corpus BREF-80), et 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.

La transcription orthographique produite de comporte aucune marque de ponctuation, aucune lettre capitale, les accords en genre et en nombres peuvent ne pas être marqués. Il peut aussi y avoir des insertions et/ou des omissions.

2.2. Le Français vers IF

Ce module [Blanchon & al. 2000, Boitet & Guilbaud 2000] 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 deux sites : à Grenoble (sur un IBM 9221-130 à 3,5 mips), au CCSJ Marseille (sur un IBM 9672-R14 à 40 mips). Lors de la démonstration du 22 juillet Ariane fonctionnait à Montpellier au centre de calcul IBM (sur un IBM 9672-RX5 à 60 mips).

Il reçoit en entrée une transcription orthographique de l'énoncé oral. Parce qu'elle peut être mal formée, on ne peut pas mettre en œuvre un processus d'analyse du type de celui que l'on met en œuvre pour faire de l'analyse d'un texte écrit qui est bien formé. De plus, l'analyseur doit être capable de faire l'analyse en composants de l'IF. Les étapes suivantes se succèdent :

- Analyse morphologique et lemmatisation des mots du texte ;
- Premières 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.

2.3. L'IF vers Français

Ce module [Wehrli & Wehrle 98] 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.

Il reçoit en entrée une IF qui est bien formée. On peut donc mettre en œuvre un processus très déterministe pour la génération. Pour cela, il faut et il suffit que l'IF reçue ait une structure couverte par le générateur. Le processus est mis en œuvre au moyen des étapes suivantes :

- 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 de GBGen pour produire une forme orthographique correctement orthographiée et ponctuée.

2.4. La synthèse du français

Le module du LAIP [Keller 97, Keller & Zelner 98] 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 généraux (7000 mots) et spécialisés (noms propres de la tâche). La génération de la prosodie met en œuvre des règles de type psycholinguistique et la génération du signal utilise la technique Mbrola de la faculté de Mons en Belgique.

3. Architecture

Les démonstrateurs des différents partenaires communiquent sur deux canaux différents, un canal pour la visioconférence et un canal pour la traduction. Localement, nous utilisons aussi le même genre d'architecture.

3.1. Intégration des démonstrateurs

En ce qui concerne la visioconférence, elle est mise en œuvre avec des systèmes commerciaux. Afin d'assurer une bonne qualité d'image pour le public. Nous avons utilisé des systèmes à 384 kilo bits par seconde.

Pour la traduction proprement dite, les échanges se font via un serveur de communication global sur lequel se connectent les sites qui participent à la conversation. La connexion au serveur de communication se fait au moyen du protocole Telnet. Un message envoyé sur le serveur de communication est reçu par tous les sites connectés qui décident ou non d'y répondre. Les données échangées sont des messages textuels, respectant une certaine syntaxe, qui contiennent des informations de trace (hypothèses de reconnaissance, générations locales), de contrôle (système prêt, ignorer l'IF précédente), et des IFs qui permettent à la traduction du dialogue de s'effectuer.

3.2. Intégration les composant du groupe CLIPS++

Tous les composants que nous avons développés sont des serveurs. Aucun d'entre eux ne communique directement avec les autres. Ils sont tous connectés à un serveur de communication local qui est donc une sorte de tableau noir. Nous avons choisi cette architecture car elle est très efficace pour le travail distribué et l'intégration de composants développés par différentes équipes. Seul le protocole d'échange est partagé.

3.3. Interface du démonstrateur du groupe CLIPS++

L'environnement de traduction n'étant pas intégré dans l'environnement de visioconférence l'utilisateur a devant lui deux écrans, celui de la visioconférence et celui de l'interface du système.

Dans la situation où l'utilisateur joue le rôle d'un client, l'écran d'interface est divisé en deux parties. Sur la partie droite, se trouve l'interface de pilotage du système de traduction proprement dit. Sur la partie gauche se trouve un environnement capable d'afficher des pages web proposée par l'agent. L'interface est montrée et décrite plus précisément en annexe 1.

Dans la situation où l'utilisateur est un agent de voyage, en dessous du serveur d'affichage de pages web se trouve un sélecteur de pages web à envoyer au client qui donnent des informations sur les transports, les hôtels, et des informations touristiques.

4. Évaluation et perspectives

Malgré de succès qu'ont rencontré nos démonstrations[3] auprès du public et les bons échos relayés par les médias, la qualité de la traduction de parole de dialogues finalisés n'a pas encore atteint une qualité et une robustesse suffisantes pour être mise sur le marché. Dans cette section, nous allons commenter notre démonstrateur et envisager des pistes qui doivent permettre d'atteindre une meilleure qualité dans le future.

4.1. Évaluation de notre démonstrateur

Nous avons mis en œuvre une architecture séquentielle dans laquelle les modules ne partagent aucune donnée et s'échangent des structures de données très simples. De ce point de vue, cette architecture n'est pas meilleure que celle qui est mise en œuvre dans l'approche "quick and dirty." Nous ne profitons pas, pour l'instant, de notre connaissance approfondie des différents modules mis en œuvre.

Aucune mémoire du passé n'est utilisée et aucun module de traitement du dialogue n'est mis en œuvre bien que nous ayons passé du temps dans la construction de modèles de dialogues.

L'IF n'est encore pas complètement couverte ni en analyse, ni en génération. À cause de sa pauvre spécification, nous nous sommes approprié son fonctionnement en utilisant d'une base de donnée d'exemples. Cela nous a fait perdre du temps et a ralenti nos développements.

Les modules français vers IF et IF vers français sont distribués. Pour la reconnaissance de la parole, un processus client tourne sur la machine du client et le serveur de reconnaissance est lui aussi distribué. Le signal parole est transmis sur le réseau local et consomme une assez large bande passante. La synthèse de la parole est elle aussi distribuée de sorte qu'elle ne se fait pas sur la machine du client.

Des progrès ne pourront être accomplis que si nous pouvons mettre en œuvre une architecture plus orientée vers le traitement du dialogue, plus intégrée, plus interactive et plus personnalisable.

---

[3] Le 22 juillet comme client avec trois agents de voyages en Allemagne, Corée du Sud et Etats-Unis et comme agent de voyage pour la démonstration d'ETRI avec le Japon et les Etats-Unis. Les 11, 13 et 14 octobre 1999 comme client de ETRI à TELECOM-99 (Genève). Le 27 octobre 1999 comme client avec la Corée du Sud et les Etats-Unis, en démonstration privée pour IBM-France.

4.2. Perspectives à court terme

Nous envisageons plusieurs pistes pour améliorer de manière significative la traduction de parole. Quelques-unes d'entre elles seront explorées dans le cadre du projet européen NESPOLE![4], les autres engendreront des résultats à plus long terme.

4.2.1. Architecture client serveur complète

Pour permettre la diffusion de la traduction de parole dans le grand public, une architecture client-serveur complète est nécessaire afin que la quantité de matériels et de logiciels spécifique soit réduite au maximum du côté de l'utilisateur. En particulier, les ressources acoustiques, linguistiques et spécifiques à la tâche qui ont de grandes tailles et sont sujettes à de fréquents changements doivent être stockées uniquement sur des serveurs partagés et non chez les utilisateurs. Ainsi, les utilisateurs bénéficieront de chaque mise à jour à la volée.

Dans le contexte de NESPOLE!, nous envisageons, dans un premier temps, d'intégrer les services de la façon suivante, en utilisant le protocole H323 pour encapsuler les données échangées entre l'agent et le client.

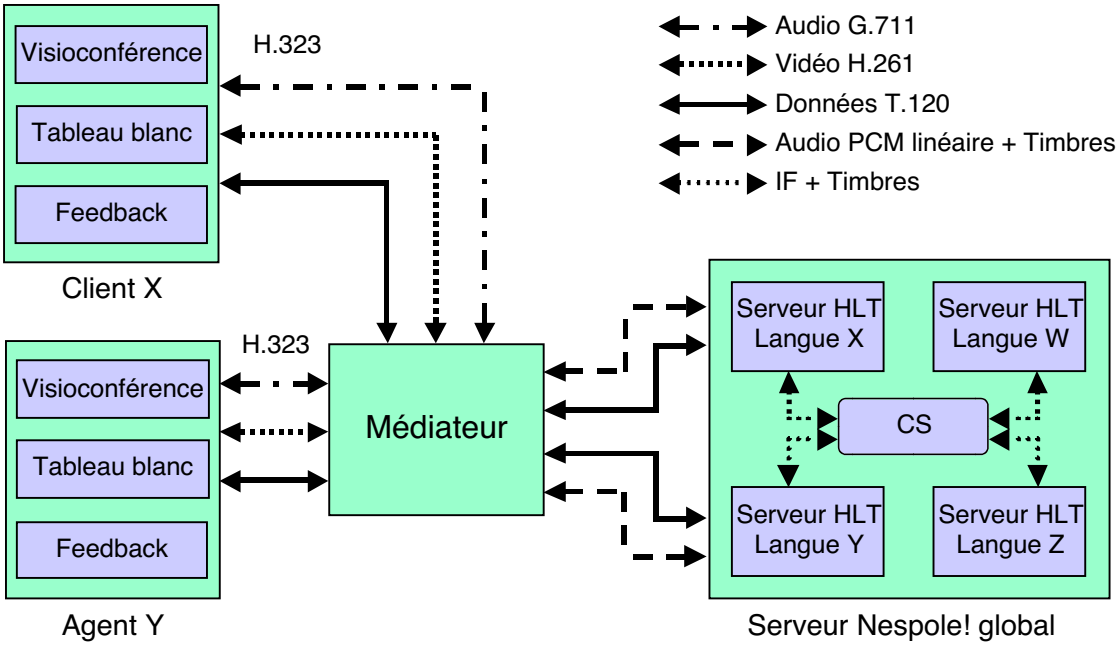


Figure 2 : Architecture du premier démonstrateur NESPOLE!

4.2.2. Traitement du contexte

Nous pensons que 3 types de contextes doivent être représentés et utilisés : le contexte global, le contexte du dialogue, et le contexte linguistique.

Le contexte global doit permettre de représenter, au moins, le type du dialogue, les caractéristiques des participants. Nous pensons en particulier à leur nom, leur sexe[5], leur age, leur niveau de politesse relative, leurs intentions – si elles sont disponibles –, et peut être leur localisation car elle peut éventuellement personnifier les intervenants[6].

---

[4] Pour plus d'informations et pour l'actualité du projet se référer à http://nespole.itc.it/

[5] En allemand ou en japonais, les noms propres doivent être utilisés lors des saluations. "Bonjour Monsieur" est possible en français, mais on ne peut pas dire "Guten Tag (mein) Herr" en allemand. Il faut dire "Gutten Tag Herr Müller". Et si un Japonais dit "Smith-san", il faut être capable de rétablire la distinction "Mr Smith", "Mme Smith", "Melle Smith".

[6] "Mais Taejon vient de me dire que …"

Le contexte du dialogue doit contenir une représentation du passé du dialogue, l'étape en cours si elle respecte un script connu, et des prédictions sur le futur. À court terme, beaucoup serait déjà possible si l'analyseur pouvait disposer d'une liste triée d'actes de parole prédis par un modèle de dialogue adéquat.

La partie la plus importante du contexte linguistique est une liste de centres possibles, soit, une liste de référents possibles pour les éléments anaphoriques et les ellipses. Voici un exemple d'exploitation possible pour un dialogue français allemand :

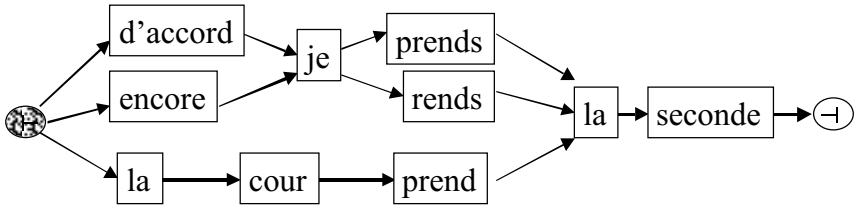
(1a) Nous avons deux chambres, une sur cour avec WC et l'autre sur rue avec douche et WC[7].
…2 Zimmer,…
(1b) Pour aller à la gare, ne prenez pas la première rue à droite, mais la seconde[8].
…die erste Straße…
(2) D'accord, je prends la seconde[9].
Einverstanden, ich werde das/die zweite nehmen.

Lorsque l'on traduit la phrase (2), le genre sera neutre dans le cas où elle succède à la phrase (1a), mais il sera féminin avec la phrase (1b). L'idée est donc d'enrichire le texte à manipuler (une chaîne orthographique en analyse et une IF en génération) avec toutes les informations disponibles.

Voici un exemple d'entrée possible pour l'analyseur [10]:

```
<ctxt_glob> <speaker> client <client> Madame Durand 70 years <agent> Herr
Biedemeyer 52 years <firme> NTG <topic> hotel reservation</ctxt_glob>
<ctxt_dial> <stage> central episode <past_sp_acts> question-info info
request <future_sp_acts> yes no question-info </ctxt_dial>
<ctxt_ling> pension-hotel_NF réserver_VT chambre_NF cour_NF réserver_VT
pension-régime_NF prendre_VT rue_NF </ctxt_ling>
<utterance> <alt> d'accord_/_encore je prends_/_rends la seconde <alt> la
cour prend la seconde </utterance>
```

Vue graphique de la reconnaissance du tour de parole `<utterance> … </utterance>` dans l'exemple ci-dessus :



4.3. Perspectives à moyen terme

4.3.1. IF mieux définie et plus riche

Le besoin d'une IF mieux structurée avec une spécification claire et un meilleur pouvoir d'expression a été bien compris par tous. Dans le cadre du projet NESPOLE! nous sommes en train de travailler dans cette direction.

[7] We have 2 rooms, one on the back with WC and the other on the street with shower and WC.

[8] To go to the station, don't take the first street on the right, but the second.

[9] OK, I'll take the second one.

[10] Nous anticipons ici sur la section « Meilleure intégration des composants »

4.3.2. Prosodie

Nous souhaitons aussi que le module de reconnaissance de la parole produise des marques prosodiques qui puissent être utilisées par l'analyse et codées dans l'IF.

Inversement, les marques prosodiques contenues dans l'IF pourraient être utilisées par le générateur, en plus d'autres traits sémantiques et pragmatiques (intentions de l'utilisateur), pour produire une sortie contenant des marques ou des étiquettes utilisable par la synthèse afin d'obtenir une synthèse adaptée au dialogue.

4.3.3. Interaction système-utilisateur et feedback

Du point de vue ergonomique, 2 approches complémentaires doivent être suivies :

- Adaptation du système à l'utilisateur (mise en œuvre de profils utilisateur),
- Adaptation de l'utilisateur au système (conseils pour rendre la parole plus facilement reconnaissable et traduisible).

4.4. Perspectives à plus long terme

4.4.1. Meilleure intégration des composants

L'intégration directe des composants est difficile et contredit la recherche de modularité et l'architecture client serveur. Les possibilités offertes pour une amélioration sont :

- L'utilisation de structures interfaces plus riches entre les composants, telles que des arbres de treillis ou des arbres de cartes,
- L'utilisation de ressources linguistiques communes (bases de données lexicales et grammaticales partagées pour produire les données linguistiques nécessaires à chaque module),
- La mise en œuvre d'architectures utilisant, par exemple, le pipe-line, des agents, un tableau noir, un tableau blanc.

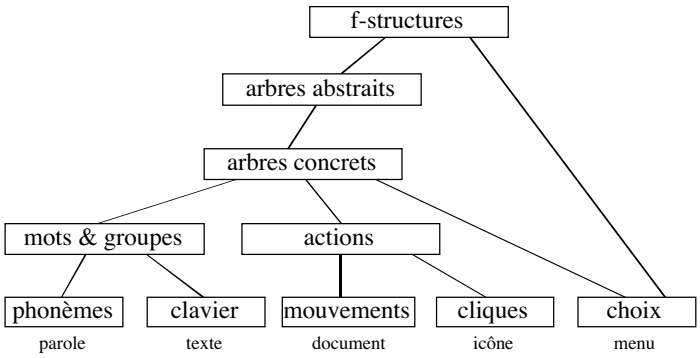


Figure 3 : hiérarchie d'intégration des modalités

4.4.2. Vraie multimodalité

Finalement, il nous semble souhaitable, bien que très difficile, de développer un système vraiment multimodal. L'une des manière d'aborder le problème est peut-être d'écrire des grammaires unimodales pour chacun des canaux et des ensembles de règles multimodales sur des couches organisées en hiérarchie comme illustré en figure 3.

5. Conclusion

Lorsque nous avons commencé à travailler en traduction de parole, nous n'étions pas vraiment certain d'atteindre une qualité de résultats suffisante, mais les progrès ont été réalisé de façon quasi exponentielle. Il nous a semblé qu'il fallait à tout prix que le français soit représenté dans ce domaine de recherche. L'effet d'entraînement de la coopération inter partenaires a aussi joué un rôle majeur.

Certes, il reste encore bien du chemin à parcourir pour atteindre la qualité requise pour la mise sur le marché de cette technologie. Nous poursuivons le travail dans le cadre du projet

NESPOLE![11] en mettant en œuvre les points de perspectives que nous avons évoqués ici.

Remerciements

Nous tenons à remercier IBM France qui nous a permis d'atteindre des temps de réponse raisonnables pour le module français vers IF. Leur intérêt pour nos travaux et la publicité qui en est faite peut être une source de débouchés intéressants.

Nous sommes aussi très reconnaissant envers l'institut I, notre université et le laboratoire C qui ont financer une grande partie de nos travaux sur ce thème.

Références

Blanchon H. & al. (1999) *Participation Francophone au Consortium C-STAR II*. La tribune des industries de la langue de du multimédia vol. **31-32** August-December 1999, pp. 15-23.

Boitet C. & Guilbaud J.-P. (2000) *Analysis into a formal task-oriented pivot without clear abstract semantic is best handled as usual translation*. Proc. ICSLP-2000.

Keller E. (1997) *Simplification of TTS architecture vs. Operational quality*. Proc. EUROSPEECH'97, September 1997, Rhodes, Greece.

Keller E. & Werner S. (1997) *Automatic Intonation Extraction and Generation for French*. Proc. 14th CALICO Annual Symposium, September 1997, West Point, NY, USA.

Keller E. & Zellner B. (1998) *Motivations for the prosodic predictive chain*. Proc. ESCA Symposium on Speech Synthesis, Jenolan Caves, Australia, vol. 1/1, pp. 137-141.

Lavie & al. (2000) *The Janus-III translation system*. To appear in Machine Translation.

Levin L. & al. (1998) *An Interlingua Based on Domaine Actions for Machine Translation of Task-Oriented Dialogues*. Proc. ICSLP'98, 30th November - 4th December 1998, Sydney, Australia, vol. 4/7, pp. 1155-1158.

Park J. & al. (1998) *Spontaneous Speech Translation System Development (in Korean)*. KSCSP vol. **15**/1, pp. 281-286.

Seligman M. & al. (1998) *Dictated Input for Broad-coverage Speech Translation*. Proc. Association for Machine Translation in the Americas (AMTA-98), Workshop on Embedded MT Systems: Design, Construction, and Evaluation of Systems with an MT Component. October 28, 1998, Langhorne, PA, USA.

Sugaya F. & al. (1999) *End-to-End Evaluation in ATR-MATRIX Speech Translation Sytem between English and Japanese*. Proc. EUROSPEECH'99, September 5-9, 1999, Budapest, Hungary, vol. 6/6, pp. 2431-2434.

Sumita E. & al. 1999) *Solutions to Problems Inherent in Spoken-language Translation: The ATR-MATRIX Approach*. Proc. MT-Summit VII, Singapore, September 13-15, 1999, vol. 1/1.

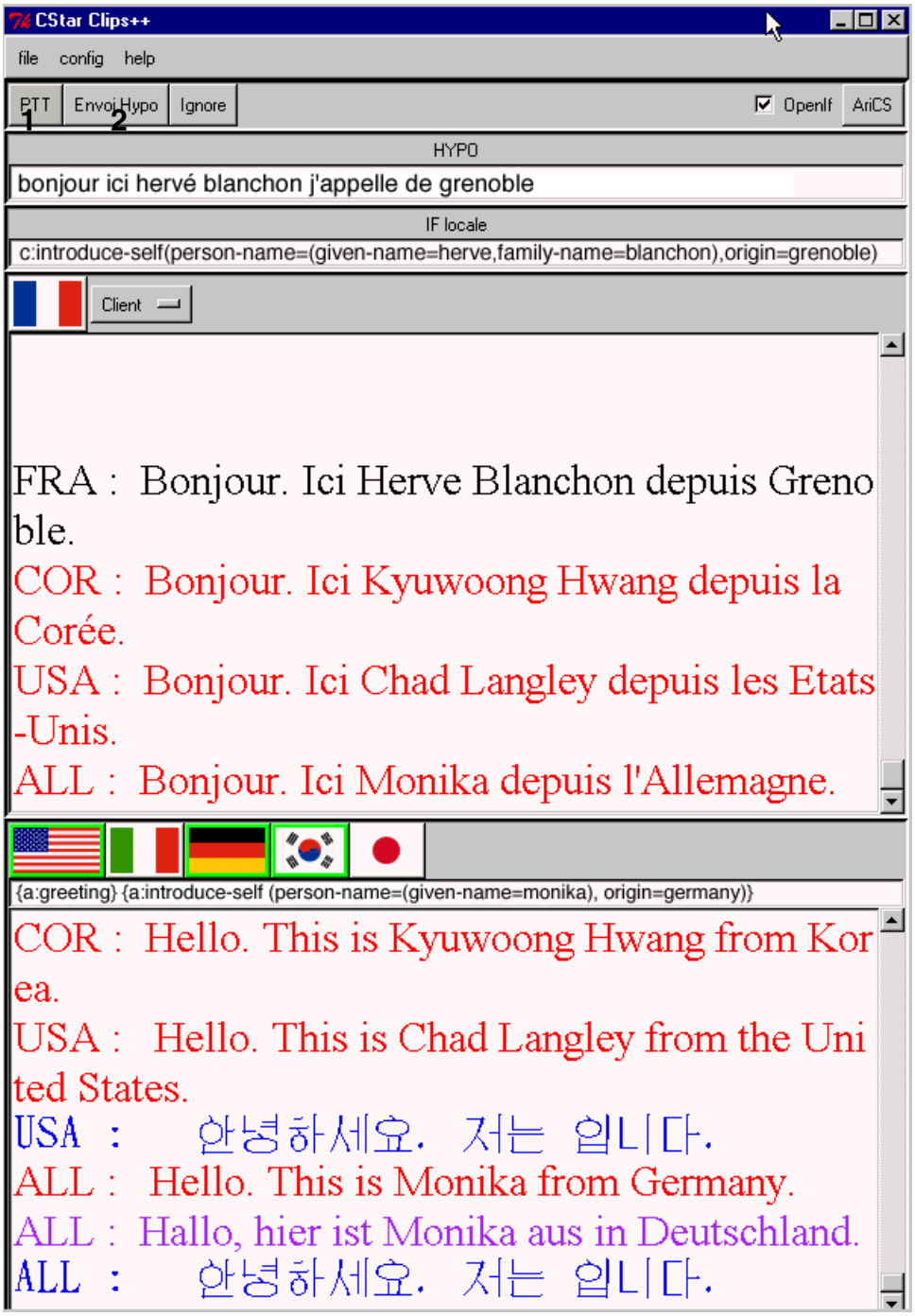
Takezawa T. & al. (1998) *A Japanese-to-English speech translation system: ATR-MATRIX*. Proc. ICSLP-98, pp. 957-960.

Vaufreydaz D. & al. (1999) *A Network Architecture for Building Application that Use Speech Recognition and/or Synthesis*. Proc. EUROSPEECH'99, Budapest, Hungary, September 5-9, 1999, vol. 5/6, pp. 2159-2162.

Wehrli E. & Wehrle E. (1998) *Overview of GBGen*. Proc. 9th International Workshop on Natural Language Generation, Niagara-on-the-lake, Canada, August 1998.

[11] Consulter http://nespole.itc.it/ pour des informations complètes et régulièrement mises à jour.

Annexe 1 : description de l'interface du démonstrateur



1) ***Mise en attente de la reconnaisse de parole***
2) ***Résultat de la reconnaissance***
3) ***Envoi de l'hypothèse de reconnaissance vers le module Français –IF***
4) ***Texte IF produit***
5) ***Rétrogénération en français pour le contrôle***
6) ***Réponse des autres participants préfixée par leur nationalité***
7) ***Langues disponibles pour l'affichage***
8) ***Dernière IF reçue (ici celle de l'Allemagne)***
9) ***Trace de la génération dans les autres langues.***