

Télécommunications et micro-informatique, les alliés du traducteur d'aujourd'hui SYSTRAN s'adapte...

Claude Bureau

Volume 37, numéro 4, décembre 1992

Études et recherches en traductique / Studies and Researches in Machine Translation

URI : <https://id.erudit.org/iderudit/002179ar>

DOI : <https://doi.org/10.7202/002179ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer cet article

Bureau, C. (1992). Télécommunications et micro-informatique, les alliés du traducteur d'aujourd'hui : SYSTRAN s'adapte... *Meta*, 37(4), 595–609.
<https://doi.org/10.7202/002179ar>

Résumé de l'article

Ces dix dernières années ont connu deux phénomènes culturels : 1) la communication, et son développement exponentiel dans tous les domaines, 2) la micro-informatique, et les PC dont la puissance s'est développée de façon inversement proportionnelle au coût. Cette évolution dans nos mœurs touche également le traducteur humain, qui peut à présent, sans s'endetter inconsidérément, se permettre d'avoir à domicile un PC performant et un MODEM. Avec ces deux atouts, le voilà propriétaire d'une fenêtre ouverte sur l'univers de l'information terminologique, et d'un outil convivial pour en tirer le meilleur parti. SYSTRAN, lui aussi, s'est adapté à cette évolution. En effet, il n'est plus ce gros système dont seuls les organismes internationaux et les multinationales pouvaient se permettre l'utilisation. À présent, n'importe qui a la possibilité d'utiliser SYSTRAN, de chez lui ou du bureau, grâce à toute une gamme d'accès peu coûteux : EXPRESS TRADUCTION, Minitel, les dialogues électroniques bilingues. Aujourd'hui, l'utilisateur de SYSTRAN peut également influencer la traduction automatique en créant son glossaire personnel « préférentiel » et « prioritaire » : SYSTRAN en tiendra compte dans le choix de la terminologie. Accessibilité et flexibilité sont devenues les mots d'ordre de la TA et de la TAO des années 90.

TÉLÉCOMMUNICATIONS ET MICRO-INFORMATIQUE, LES ALLIÉS DU TRADUCTEUR D'AUJOURD'HUI

SYSTRAN s'adapte...

CLAUDE BUREAU
SYSTRAN S.A., Paris, France

Résumé

Ces dix dernières années ont connu deux phénomènes culturels :

- 1) la communication, et son développement exponentiel dans tous les domaines,*
- 2) la micro-informatique, et les PC dont la puissance s'est développée de façon inversement proportionnelle au coût.*

Cette évolution dans nos mœurs touche également le traducteur humain, qui peut à présent, sans s'endetter inconsidérément, se permettre d'avoir à domicile un PC performant et un MODEM.

Avec ces deux atouts, le voilà propriétaire d'une fenêtre ouverte sur l'univers de l'information terminologique, et d'un outil convivial pour en tirer le meilleur parti.

SYSTRAN, lui aussi, s'est adapté à cette évolution.

En effet, il n'est plus ce gros système dont seuls les organismes internationaux et les multinationales pouvaient se permettre l'utilisation.

À présent, n'importe qui a la possibilité d'utiliser SYSTRAN, de chez lui ou du bureau, grâce à toute une gamme d'accès peu coûteux :

EXPRESS TRADUCTION

Minitel

les dialogues électroniques bilingues.

Aujourd'hui, l'utilisateur de SYSTRAN peut également influencer la traduction automatique en créant son glossaire personnel «préférentiel» et «prioritaire» : SYSTRAN en tiendra compte dans le choix de la terminologie.

Accessibilité et flexibilité sont devenues les mots d'ordre de la TA et de la TAO des années 90.

Abstract

These past ten years have witnessed two cultural phenomena:

- 1) communication and its exponential development in every domain,*
- 2) micro-computing and the PCs, whose increase in power has been inversely proportional to their cost.*

This evolution in our everyday life has also had an impact on the human translator, who can now afford, without going too much into debt, to have a high performance PC and a modem at home. These two give him both a window opening onto the world of terminological information and a user-friendly tool to make the most of it.

SYSTRAN too has adapted to this evolution. Instead, it is no longer that huge system which only international organizations and multinational concerns could afford to use.

Nowadays, anyone can use SYSTRAN, from home or the office, thanks to a whole series of inexpensive access possibilities:

*EXPRESS TRANSLATION**Minitel services**electronic bilingual dialogues**Today, SYSTRAN users can also influence machine translation by creating their own personal «preference» and «default» glossaries: SYSTRAN will take them into account during lexical choice.**Accessibility and flexibility have become the imperatives MT and MAT must meet in the '90s.***1. SYSTRAN ÉVOLUE****1.1. LE SYSTÈME**

SYSTRAN est le premier système de traduction automatique. Dans les années 1965/1985, il est passé par une étape qui a vu la dispersion de ses moyens et de ses droits. Chaque couple de langues constituait un bloc indissociable : pour traduire entre elles les 9 langues de la Communauté Européenne, il était nécessaire de développer 72 systèmes.

En 1986, la Société GACHOT a racheté les Sociétés créatrices WTC et LATSEC aux États-Unis, ainsi que SYSTRAN INSTITUT en Allemagne et SYSTRAN LUXEMBOURG. En regroupant cet ensemble sous la coupe de sa filiale SYSTRAN S.A., elle a formé une équipe unique qui a permis de restructurer le système sur de nouvelles bases. Celui-ci a retrouvé des possibilités de développement qui s'étaient perdues dans la diversité des besoins de ses clients multiples.

Les programmes sont découpés en trois ensembles : Analyse, Transfert et Synthèse. Chacun regroupe des modules indépendants réalisant une tâche bien déterminée.

1) L'analyse

C'est la partie la plus importante du programme qui représente pour un couple de langues pratiquement les deux tiers des 130 000 lignes de programmation.

Son objectif est d'identifier la structure de la phrase en langue source, en traitant lors de passes successives les divers types possibles d'interrelations syntaxiques les uns après les autres.

À chaque étape, elle prend des décisions sur l'existence ou la non-existence de ces interrelations entre les différents mots :

- levée des ambiguïtés homographiques,
- découpage des phrases en proposition principale et subordonnée,
- détermination des relations syntagmatiques de base,
- établissement des énumérations,
- détermination du prédicat principal,
- recherche du sujet et des antécédences,
- analyse des relations entre des propositions,
- recherche du gouvernement des prépositions.

2) La synthèse

Cet ensemble est chargé de restituer, selon le génie de la langue cible, ce qui a été analysé dans la langue source ; il est constitué :

- d'un module de synthèse proprement dite qui détermine le genre, le cas, le nombre, le temps des différents termes,
- d'un programme de réarrangement syntaxique,
- et d'un programme de génération morphologique.

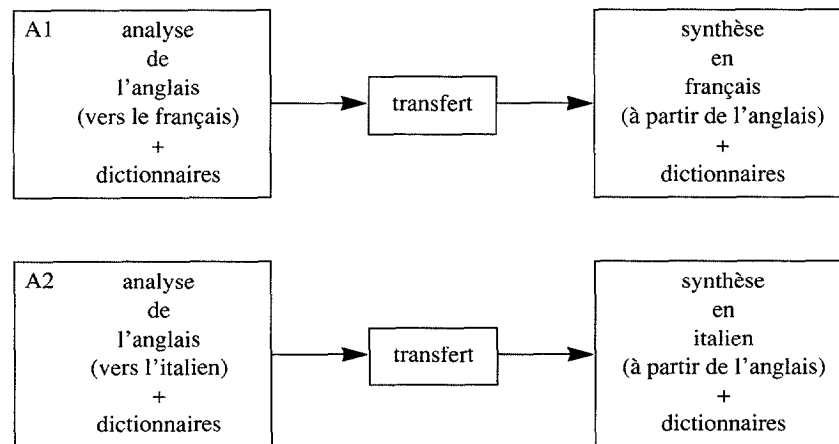
3) Le transfert

Entre l'analyse et la synthèse, il existe une phase très importante, dite de transfert, s'appuyant particulièrement sur le dictionnaire contextuel dont nous parlerons ci-après. Composée de routines lexicales permettant le passage de la structure d'une langue à celle d'une autre, elle est donc particulière au couple «langue source — langue cible».

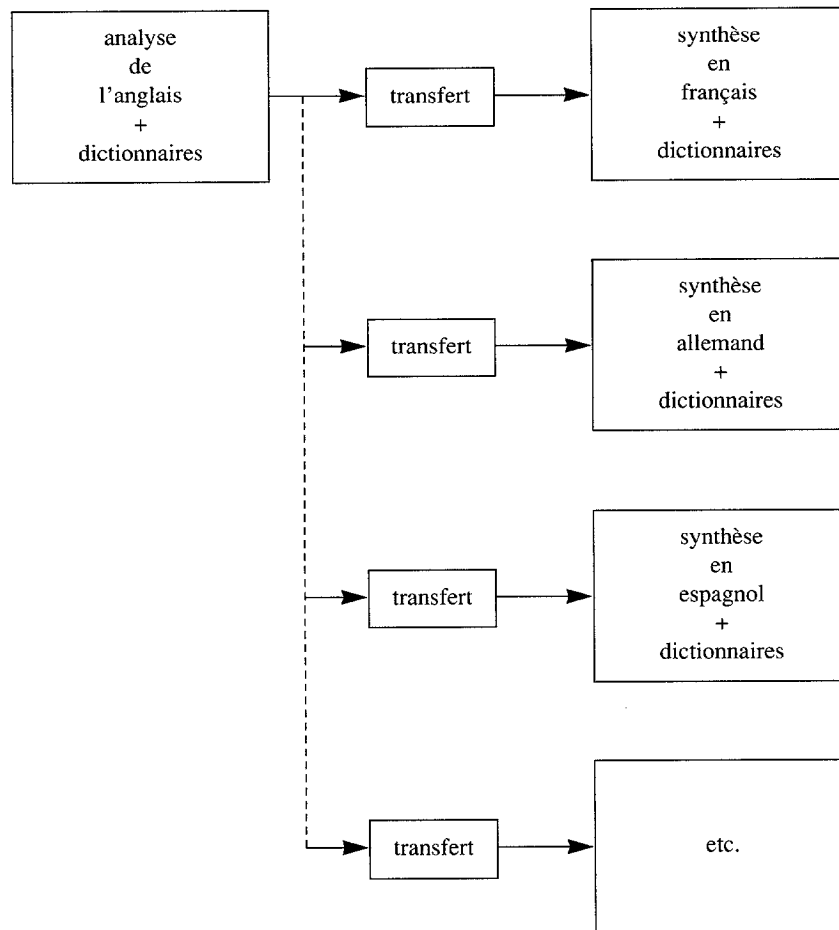
L'un des principaux buts de la modernisation du SYSTRAN a été de rendre le système modulaire et multicible, c'est-à-dire qu'à partir de l'analyse d'une langue source on peut développer plusieurs paires de langues, par exemple, l'anglais vers le français, l'anglais vers l'allemand, l'anglais vers l'espagnol. Il suffit alors d'écrire les 9 analyses de langue source, les 9 synthèses de langue cible ; cependant les 72 programmes de transfert restent toujours nécessaires, car ils représentent les «prises multiples» permettant de décliner une nouvelle paire de langues à partir d'autres déjà existantes.

À ce jour, 23 paires de langues ont été développées à des niveaux de qualité divers. Quinze paires sont opérationnelles et utilisées professionnellement (*cf.* section 2.2.).

Système monocible



Système multicable



Le SYSTRAN constitue ainsi un ensemble très élaboré et très structuré de programmes et de dictionnaires qui va réaliser une série de va-et-vient sur les mots de chaque phrase, jusqu'à définir l'ensemble des interrelations existant entre tous ces mots.

Il se base, pour ce faire, sur des renseignements concernant les associations possibles entre les mots ou les caractéristiques grammaticales, syntaxiques, sémantiques et prépositionnelles de chaque mot qui lui sont données par les dictionnaires de base.

Après l'analyse, un choix est réalisé entre les multiples sens possibles des expressions grâce à un dictionnaire spécial dit des «expressions à cadre sémantique limité». Celui-ci est, en fait, une base de données de micro-programmes qui vont décider du sens adéquat de chaque mot en fonction du contexte syntaxique et sémantique.

Une fois la phase de transfert terminée, le programme de synthèse va rédiger à nouveau le texte dans la langue cible en appliquant l'ensemble des règles grammaticales et syntaxiques programmées.

Les étapes parcourues lors de la traduction se présentent comme suit :

- pré-processing,
- expressions idiomatiques invariables,
- dictionnaire principal syntaxique et sémantique,
- dictionnaire des groupes nominaux,
- levée des ambiguïtés homographiques,
- relations entre propositions élémentaires,
- groupes syntagmatiques,
- relations analytiques,
- recherche des énumérations,
- dictionnaires analytiques,
- recherche du sujet,
- relations intrinsèques,
- dictionnaires des expressions verbales et conditionnelles,
- traduction des prépositions,
- structures spéciales,
- régénération grammaticale,
- réarrangement de l'ordre des mots,
- post-processing.

1.2. LES DICTIONNAIRES

Quelle que soit la complexité des programmes linguistiques qui représentent 130 000 lignes de programmation en moyenne par paire de langues, leur importance est relativement faible par rapport à celle des dictionnaires eux-mêmes.

En effet, le coût des dictionnaires représente près de 90 % du coût de développement d'une paire de langues considérée. S'il est en effet facile de développer des dictionnaires de 20 000 ou 30 000 termes, leur complexité s'accroît très vite avec l'augmentation de la taille, car les descriptions techniques générales ne sont plus suffisantes pour distinguer les différents sens d'un mot donné dans un domaine particulier.

Un dictionnaire classique (papier ou électronique) va préciser les points suivants pour un nom : son orthographe, le genre, le nombre, sa traduction... ce qui représente 5 ou 6 «codes», alors que l'encodage complet dans le dictionnaire du SYSTRAN comporte jusqu'à 165 positions qui permettent de définir le mot ou l'expression le plus complètement possible, afin que le système puisse prendre les options nécessaires le moment voulu.

Les dictionnaires peuvent être classés comme suit :

- 1) Le dictionnaire des mots de base :
il contient pour chaque mot de la langue source une description de celui-ci à la fois morphologique, grammaticale, syntaxique et sémantique, complétée par la traduction de base dans chacune des langues cibles accompagnée par l'ensemble des indications grammaticales et syntaxiques correspondant à la synthèse.

Dans ce dictionnaire, les différents homographes sont traités en tant que substantif, verbe ou adjectif... cependant les polysémies ne sont pas traitées à ce niveau, mais relèvent des dictionnaires contextuels.

2) Les dictionnaires d'expressions :
c'est à ce niveau que les mots seront traités en fonction du contexte dans lequel ils se trouvent, en répartissant les dictionnaires selon la liste ci-après :

- a) **Dictionnaire idiomatique** dans lequel seront classées les expressions invariables, par exemple «à la queue leu leu».
- b) **Dictionnaire des groupes nominaux** : il s'agit là aussi d'expressions figées, mais variables en nombre, du type «*pomme de terre*», «*chemin de fer*», ou «*ville d'eau*».
- c) **Dictionnaires analytiques** : ils traitent des exceptions aux règles grammaticales stockées dans les programmes. Un sous-ensemble, les dictionnaires de résolution homographique, permet de prendre en charge certains termes homographiques en fonction d'un contexte très précis, du type «*rendre compte*», «*tenir compte*», ou «*prendre pied*».
- d) **Dictionnaire contextuel** : celui-ci est de loin le plus complexe, car c'est lui qui doit permettre de faire le choix final entre les différents sens d'un terme. Ainsi le mot «*oil*» en anglais nécessite le codage de plus de 400 contextes pour pouvoir déterminer le sens à retenir parmi les trois possibles : «*huile*» (mécanique ou alimentaire), «*pétrole*», ou «*parfum*». Le terme «dictionnaire» est d'ailleurs ici mal choisi, car il s'agit en fait d'une base de données de micro-programmes linguistiques (syntactiques et sémantiques) exécutés sur la base d'un mot clé source.

Ainsi peuvent être prises en considération des différences telles que :

- «*Il fait beau*» ou «*il a beau faire*».
- Prendre «*un*» verre ou prendre «*le*» verre.

Ces dictionnaires constituent un système expert organisé dans une base de données en dix niveaux, subdivisés chacun en huit niveaux de priorité. Le tout est géré par un véritable moteur d'inférence indépendant de la langue source et de la langue cible, et qui définit à chaque étape l'applicabilité de chaque règle lexicale.

1.2.1. LE DÉVELOPPEMENT SÉMANTIQUE

Les nécessités de l'automatisation obligent à expliciter toutes les informations. Cette «mise à plat» des données linguistiques ne concerne pas seulement la morphologie (codage systématique des flexions, des variantes graphiques, des homographes); elle porte directement sur le codage sémantique.

Par exemple, pour le mot anglais «*Vessel*», un codage spécifique dans le dictionnaire contextuel permet, s'il est sujet d'un verbe de mouvement, de le traduire par «*navire*» et non pas par «*réceptient*». Cette disposition permet à SYSTRAN de choisir entre les deux sens possibles de ce mot, en fonction du verbe utilisé dans la phrase source.

La rigidité et la lourdeur d'un arbre sémantique, mettant en jeu quelque 500 catégories selon une hiérarchie très contraignante, a conduit SYSTRAN à développer la notion de «classes d'objets» (profession, vêtement, voie de communication...) pour définir plus simplement les interrelations entre les mots ou expressions.

Contrairement aux catégories antérieures, les classes d'objets ne sont pas des entités logiques élaborées *in abstracto*, mais sont en prise directe avec la fonction exercée.

Les codes sémantiques affectés aux «mots de base» permettront de donner une signification plus large aux codifications qui se trouvent actuellement dans le dictionnaire contextuel.

Ceci permet de traiter plus simplement les polysémies, notamment dans le cas de verbes tels que «prendre» ou «porter» :

<i>prendre</i>	un <moyen de transport>	= <i>to take</i>
	une <maladie>	= <i>to catch</i>
	des <aliments>	= <i>to take</i>
<i>porter</i>	un <vêtement>	= <i>to wear</i> , au lieu de <i>to carry</i>

Les travaux effectués en ce sens permettent parallèlement de compléter la réversibilité des dictionnaires quel que soit le couple de langues concerné.

L'utilisation systématique de la notion de classes d'objets constitue un changement important dans la conception du système SYSTRAN. Elle met à sa disposition un outil très puissant pour la désambiguïsation et le traitement des polysémies, ce qui se traduira dans un proche avenir par une amélioration sensible de la qualité des traductions.

Elle est liée à l'introduction d'une récursivité limitée permettant au système de revenir sur un choix effectué lors d'une des premières étapes de l'analyse, si celui-ci peut être remis en cause ultérieurement par les informations relatives aux classes d'objets.

1.2.2. PRISE EN COMPTE DE LA TERMINOLOGIE DE L'UTILISATEUR

Outre les dictionnaires généraux (mots simples ou expressions) SYSTRAN dispose d'une série de 24 dictionnaires spécialisés dits *TG (Topical Glossaries)* dans les domaines suivants :

- agronomie,
- arts,
- aviation et espace,
- biologie et médecine,
- chimie,
- documents (Conseil de l'Europe),
- électronique,
- énergie,
- finance et économie,
- géologie,
- gestion du matériel,
- industrie des mines,
- informatique,
- ingénierie industrielle,
- juridique,
- mathématiques,
- métallurgie,
- minutes (CEE),
- nucléaire,
- optique et photographie,
- physique,
- sciences militaires,
- statistiques,
- typographie et publications.

L'utilisateur peut en choisir trois, en précisant l'ordre de priorité qu'il veut mettre en œuvre. En cas d'ambiguïté dans le choix du sens d'un terme, le système retiendra celui

précisé par le TG sélectionné en premier; en cas d'absence du terme dans celui-ci, il passera au second, puis au troisième; enfin au dictionnaire général.

Les dictionnaires spécialisés permettent donc pour les textes techniques de résoudre les polysémies en fonction du sujet traité.

En outre, deux dispositifs complémentaires permettent de tenir compte, de manière plus précise, des besoins spécifiques des utilisateurs.

Certains grands groupes utilisent une terminologie qui leur est propre du fait de leur domaine d'activité, ou bien par rapport à leurs confrères ou au produit concerné. SYSTRAN a créé un niveau de dictionnaire client dit *CUST (customer)*, qui permet de rassembler les traductions retenues pour chacun de ses clients importants: Aérospatiale, Bull, EDF, Rhône-Poulenc...

Par ailleurs l'utilisateur, au moment de lancer sa traduction, peut préciser sa terminologie personnelle en fonction de ses besoins propres ou des spécificités du texte qu'il envoie à la traduction, en la consignant dans le *CSD (Customer Specific Dictionary)*

À l'heure actuelle, les codifications qui peuvent être ainsi imposées au SYSTRAN dans le CSD portent sur:

- les noms et les adjectifs, (y compris les gérondifs et les participes passés à usage exclusivement nominal ou adjectival),
- les abréviations (ou les acronymes),
- les mots composés et locutions nominales,
- les expressions nominales.
- Les verbes pourront être codés d'ici peu.

Ce dictionnaire permet d'ajouter des termes qui n'existeraient pas dans les dictionnaires, mais aussi de modifier des entrées existantes, pour lesquelles l'utilisateur souhaite d'autres traductions que celles proposées par SYSTRAN.

Il peut être complété et mis à jour à tout moment en fonction des besoins, et même à l'occasion de la traduction d'un texte spécifique nécessitant une terminologie particulière.

Le codage est largement facilité par le menu déroulant qui permet de déterminer les codes essentiels nécessaires au SYSTRAN, sans avoir à faire un apprentissage linguistique poussé.

Ces dictionnaires peuvent contenir 1 000 entrées par couple de langues. Ils sont disponibles en anglais-français, anglais-italien, anglais-espagnol, anglais-portugais et français-anglais, et seront très prochainement étendus à d'autres couples de langues notamment à partir de/ou vers l'allemand.

L'ordre de préséance entre les différents dictionnaires retenu par le système pour résoudre les ambiguïtés lexicales sera alors le suivant:

1. CSD
2. CUST
3. TG 1
4. TG 2
5. TG 3
6. Général
7. les autres TG non sélectionnés

1.3. LES AIDES AUX TRADUCTEURS

Que les systèmes de traduction automatique ou assistée par ordinateur soient accessibles sur un serveur propre à l'entreprise, en service bureau, ou bien par réseau Minitel,

leur utilisation doit s'inscrire dans le cadre beaucoup plus large du traitement de la langue naturelle.

Pour ce faire, SYSTRAN étudie des stations de travail, notamment avec le concours de XEROX, ou dans le cadre de projets européens.

Il développe par ailleurs différents logiciels afin de faciliter la tâche aussi bien des traducteurs, des rédacteurs ou des réviseurs de traduction automatique.

Ces utilitaires sont multiples et peuvent concerner plusieurs types de tâches :

- Prétraitement :
 - récupération du texte,
 - correction orthographique,
 - préparation (marquage des mots à ne pas traduire, terminologie spécifique : CSD*).
- Envoi à la traduction :
 - accès au réseau de transmission,
 - envoi, récupération du texte,
 - conservation du format*.
- Postédition :
 - révision du texte,
 - vérification de la terminologie* :
- consultation de lexiques spécialisés,
- dictionnaires de synonymes,
- ...

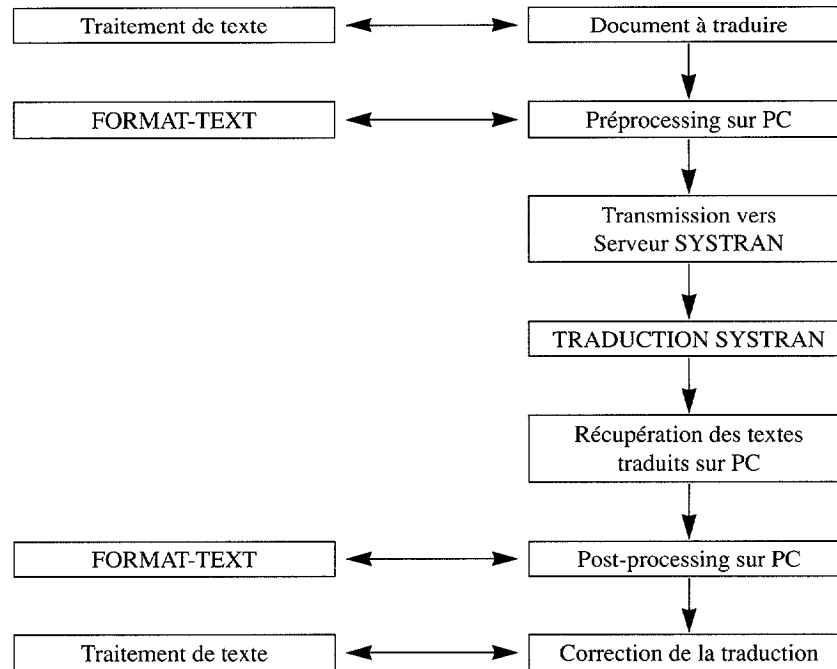
Nous n'illustrons dans cette liste que trois exemples (marqués d'un *) montrant comment SYSTRAN a pu traiter quelques-uns de ces points en gardant toujours en vue le service complémentaire à assurer, et la convivialité.

L'usage de la télématique étant de plus en plus répandu, la combinaison des nouvelles possibilités telles que micro-informatique, Minitel, réseaux à grande capacité, consultation de bases de données «en ligne» ou sur CD-ROM va permettre aux utilisateurs de trouver à leur disposition les outils indispensables à un usage professionnel.

1.3.1. CONSERVATION DU FORMAT DES TRAITEMENTS DE TEXTES

Les traductions automatiques sont effectuées à partir de fichiers de textes en ASCII, à partir des divers traitements de textes utilisés de par le monde. SYSTRAN a développé le logiciel FORMAT-TEXT qui permet de récupérer le fichier de l'utilisateur, de transmettre au serveur le seul texte en code ASCII débarrassé de ses codes de formatage, de récupérer la traduction fournie par le serveur, et de réinstaller ces codes de format avant de restituer le tout à l'utilisateur.

Ainsi le format original du texte est intégralement respecté. Si un mot du texte source figure par exemple en gras et souligné, sa traduction figurera elle-même en gras et souligné. De plus toutes les indications de mise en page, justification, marges... sont conservées.



Le logiciel utilitaire FORMAT-TEXT est actuellement pleinement opérationnel avec les traitements de textes WordPerfect et MS Word, la conversion d'autres formats est en cours de développement.

1.3.2. CONSULTATION DE DICTIONNAIRES ÉLECTRONIQUES

Lors de la rédaction, de la traduction ou de la révision d'un texte, les utilisateurs peuvent désirer consulter divers dictionnaires monolingues (synonymes, homonymes...), ou multilingues (généraux ou de spécialités).

Les dictionnaires de traduction automatique sont mal adaptés à cet usage, dans la mesure où ils possèdent une structure complexe. Celle-ci nécessite des codages extrêmement sophistiqués, rendus nécessaires par le traitement informatique, alors que le traducteur humain dispose d'éléments de jugement personnel beaucoup plus étendus.

En fait l'utilisateur a besoin de vérifier la terminologie qui lui est proposée par le traducteur automatique, mais aussi de pouvoir sélectionner la traduction appropriée d'un terme en fonction de ses propres critères et comparer les traductions proposées par les différents auteurs ou les différents ouvrages terminologiques.

SYSTRAN a donc constitué une base de données terminologiques multilingues appelée MULTIDIC à la demande et avec l'aide du ministère de la Recherche, pour permettre la diffusion la plus large possible des terminologies existantes. Il s'agit de présenter à l'écran l'image électronique des dictionnaires papier dont ils sont issus, en

indiquant à chaque fois très clairement le détenteur de cette terminologie (auteur, ouvrage...), et en passant instantanément de l'un à l'autre.

Cet ensemble de logiciels d'aide s'intègre parmi toutes les mesures d'adaptation de SYSTRAN à l'évolution de la communication et de la micro-informatique, qui permettent un accès possible par tous les moyens mis à la disposition des utilisateurs de la télématique.

2. SYSTRAN ET LA TÉLÉMATIQUE

2.1. LES SERVEURS LOCAUX : TRADUCTION AUTOMATIQUE SUR SITE CLIENT

Jusqu'en 1986, SYSTRAN ne pouvait être installé que sur de gros systèmes, et était donc réservé à des utilisateurs possédant des moyens informatiques puissants, tels que US AIR FORCE, XEROX, GENERAL MOTORS, COMMISSION DES COMMUNAUTÉS EUROPÉENNES, RHÔNE-POULENC...

Depuis quelques années la puissance des unités n'a cessé de croître, et leur coût de baisser. Des serveurs dédiés à la traduction automatique peuvent à présent être installés chez les clients sur des mini-ordinateurs interrogés par plusieurs postes de travail.

Une telle solution, qui allie les qualités de souplesse d'utilisation et de confidentialité, vient d'être mise au point à partir d'unités centrales IBM de la série 9370 ou XEROX RBS reliées à des postes de travail de traduction, telle la station XEROX 6520 (DOCUTRAN), ou des PC compatibles.

Cependant, ces serveurs étant totalement autonomes, une mise à jour périodique des dictionnaires et des programmes doit être prévue dans le cadre d'un contrat de maintenance. Dans les mois à venir, la portabilité du système sera assurée sous UNIX, permettant ainsi une plus large diffusion.

2.2. LES RÉSEAUX : TRADUCTION AUTOMATIQUE «EN LIGNE»

Depuis 1986, SYSTRAN est accessible en «service bureau» à partir des deux serveurs installés :

- l'un à Soisy-sous-Montmorency, en France, dans la banlieue de Paris,
- l'autre à San Diego, en Californie.

De n'importe quel point du monde relié à un réseau de télécommunications (TRANSPAC, INFONET, réseau national...), un utilisateur disposant d'un PC et d'un modem peut envoyer, à l'aide du logiciel de communication EXPRESS-TRADUCTION, ses textes à l'un de ces deux serveurs et obtenir les traductions en retour dans les minutes qui suivent.

L'équipement nécessaire chez l'utilisateur est le suivant :

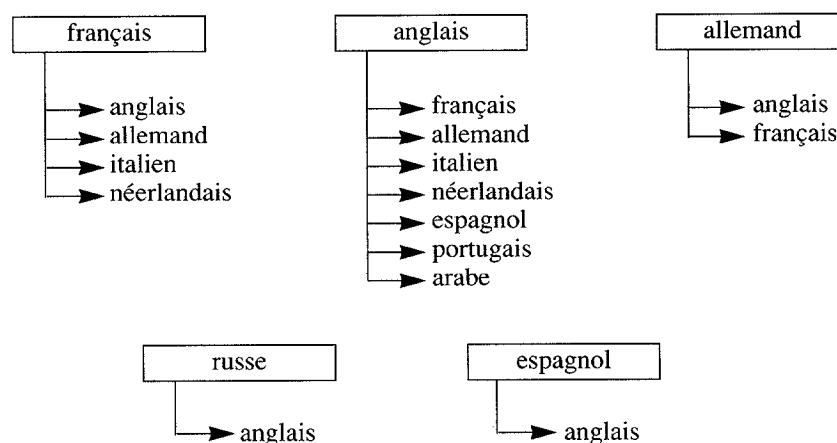
- **Matériel :**
 - IBM PC ou compatible (Min. 640 Ko RAM) ;
 - modem asynchrone, compatible Hayes, synchrone ;
 - ligne téléphonique (RTC, LS ou RNIS).
- **Logiciels PC :**
 - DOS 3/0 et plus ;
 - EXPRESS TRADUCTION ;
 - traitement de textes (WordPerfect 4.2 et plus, WORD...);
 - correcteur orthographique (optionnel) ;
 - lecteur optique (optionnel).

Le raccordement direct sur ces serveurs de réseaux internes de grandes entreprises a été réalisé. En parallèle, des essais sont en cours pour le raccordement de bornes en libre service conçues par la société KIS, et qui peuvent être installées dans les administrations, les bibliothèques, les centres d'affaires, ou tout lieu public.

Le déroulement des opérations s'effectue alors comme suit :

- Saisie et contrôle orthographique :
si le texte n'est pas déjà disponible dans la mémoire du PC ou sur disquette, la saisie se fait au clavier ou par lecteur optique.
Le texte ne doit comporter aucune faute d'orthographe, d'accentuation ou de ponctuation ! L'utilisation du correcteur orthographique permet de faciliter leur élimination.
- Langue cible :
il faut ensuite choisir la ou les «langues cibles», c'est-à-dire les langues dans lesquelles le texte sera traduit.

SYSTRAN propose actuellement sur PC 15 couples de langues :



- Dictionnaires spécialisés :
si le texte source traite d'un domaine bien particulier, on peut présélectionner 3 des 24 dictionnaires spécialisés de SYSTRAN (cf. section 1.2., la liste des TG). SYSTRAN les utilisera en priorité aux côtés du dictionnaire général ; ainsi «*hardware*» se traduira bien par «*matériel*» au lieu de «*quincaillerie*», si l'on sélectionne le dictionnaire «informatique». Le dictionnaire personnel de l'utilisateur peut être pris en compte s'il a été codé dans le CSD.
- Envoi :
le serveur reçoit le texte, le traduit et le renvoie à l'utilisateur.
Les deux textes — texte à traduire, texte traduit — s'affichent sur l'écran dans deux fenêtres disposées l'une au-dessus de l'autre. Le défilement coordonné de ces deux fenêtres permet de réviser directement le texte de la traduction électronique.

2.3. LE MINITEL

Le développement en France du Minitel, qui dispose de plus de 6 000 000 de postes répartis dans le public et les entreprises, est un phénomène en soi, mais la contagion gagne un certain nombre de pays.

Le poste Minitel est un terminal mis à la disposition des abonnés au téléphone. Muni d'un écran et d'un clavier simplifié, il permet de consulter des bases de données, ou de dialoguer dans le cadre de messagerie. Ces services sont en général accessibles sans abonnement préalable, la rémunération étant prélevée directement sur la facture du téléphone. L'accès se fait à partir du réseau téléphonique en composant un numéro de 4 chiffres correspondant au tarif appliqué, suivi d'un nom (ou de 4 autres chiffres) désignant le service.

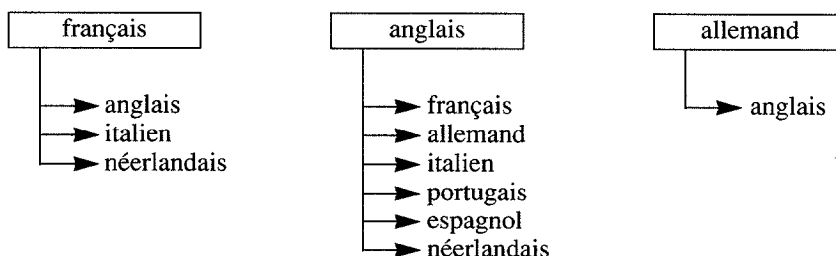
Avec le code «3617 SYSTRAN» est proposé un véritable «dispositif d'aide à la traduction» :

- le service de **traduction automatique** est destiné aux textes techniques généraux ;
- une bibliothèque personnelle permet de sauvegarder pendant deux semaines les textes traduits, à traduire ou retraduire ;
- la traduction présentée dans le cadre d'une exposition à la Cité des Sciences de la Villette (Paris) ;
- le **dialogue électronique multilingue** où deux utilisateurs situés l'un à Londres, l'autre à Barcelone par exemple, échangent des messages par l'intermédiaire du serveur SYSTRAN, chacun rédigeant ses textes et recevant ceux de ses interlocuteurs dans sa propre langue.
- une bibliothèque de **dictionnaires électroniques** est offerte à la consultation pour une recherche terminologique ou pour la révision de la traduction automatique. Deux modes d'accès sont possibles : ouvrage par ouvrage, ou directement à l'ensemble de la base de données (cf. section 2.4. ; Multidic)

L'accès au service SYSTRAN sur Minitel se fait avec des menus en français, anglais ou allemand, au choix de l'utilisateur.

Le texte à traduire est tapé sur le clavier du Minitel au cours de la consultation. Certains utilisateurs disposent cependant de microordinateurs permettant une émulation du réseau Minitel. Ceci leur permet de récupérer des textes sources préenregistrés sur fichier, de les envoyer directement à la traduction et de récupérer le texte traduit dans leurs fichiers.

Les couples de langues disponibles sur le Minitel sont au nombre de dix :



Il a été réalisé au cours de l'année dernière plus de 300 000 traductions et plus de 50 000 consultations de dictionnaires.

2.4. MULTIDIC

Sur le Minitel «3617 SYSTRAN», l'accès à la consultation des dictionnaires électroniques est analogue à celui dont disposerait un lecteur dans une bibliothèque : dans le menu d'accès, l'utilisateur choisit l'ouvrage qu'il veut consulter, précise le couple de langues dans lequel il veut obtenir la traduction de son terme, puis l'expression ou le mot recherché.

Ce mode de consultation atteint cependant très vite ses limites : la personne qui consulte peut ne pas trouver exactement la traduction qu'elle recherche dans l'ouvrage consulté, ou même ne pas trouver du tout le terme dans cet ouvrage. Elle est alors obligée de sortir de l'ouvrage, comme si dans une bibliothèque elle le reposait sur le rayon, d'en sélectionner un autre et de recommencer.

Le service Multidic permet de consulter l'ensemble de la base en une seule opération : d'entrée de jeu l'utilisateur précise le couple de langues dans lequel il veut travailler et l'expression recherchée. Lui sont alors présentés les différents ouvrages qui proposent une traduction de cette expression, et il peut alors consulter successivement chacun d'entre eux, simplement en utilisant une touche de fonction <suite> ou <retour>.

À l'heure actuelle, une vingtaine d'ouvrages peuvent être consultés sur cette base de données qui sera portée, dans les mois qui viennent, à une cinquantaine d'ouvrages.

Le tableau ci-après donne la liste des dictionnaires proposés, avec les langues accessibles.

Sauf exception, pour un ouvrage déterminé, tous les couples correspondant aux langues indiquées sont accessibles.

Multidic est actuellement consultable par le Minitel, sans abonnement, avec le code d'accès 38 28 00 28.

Cette base, qui comporte aujourd'hui essentiellement des ouvrages multilingues, sera classée en trois parties :

- lexiques ou dictionnaires multilingues,
- dictionnaires monolingues (synonymes...),
- thésauri.

Dans un très proche avenir cette base de données sera consultable en ligne, avec abonnement préalable.

L'objectif est ensuite de diffuser cette terminologie sur CD-ROM, chaque CD comportant quelques dizaines d'ouvrages réunis selon un thème principal (transport, médical, agriculture, informatique...). En plus des dictionnaires spécialisés correspondant à ce thème, d'autres ouvrages plus généraux ou connexes (juridique, affaires, synonymes...) permettront de mettre à la disposition des rédacteurs ou des traducteurs d'une profession un outil très performant.

BIBLIOGRAPHIE

De nombreux articles de presse ou des études réalisées par des groupes internationaux ont eu pour sujet la traduction automatique et le système SYSTRAN. Il est quasiment impossible de les répertorier ici. Nous nous bornerons à rappeler l'ouvrage ci-après.

HUTCHINS, W. J. (1986) : *Machine Translation. Past, Present, Future*, New York, Ellis Horwood, 382 p.

Nom de l'ouvrage	Éditeur Auteur	Langues					
		Fran.	Angl.	All.	Esp.	Ital.	Port.
DÈS AUJOURD'HUI :							
Généraux :							
Harrap Shorter		*	*	—	—	—	—
Softex		*	*	*	—	—	—
Spécialités :							
Aéronautique / Spatial		*	*	*	*	—	—
Agriculture	Mus. Hist. Nat.	*	*	—	—	—	—
Aviation	GRAISSIN	*	*	—	—	—	—
Électronique	Maison du Dict.	*	*	—	—	—	—
E.A.O.	DON (Ophrys)	*	*	—	—	—	—
Homonymes français	J. CAMION	*	*	—	—	—	—
Informatique	Maison du Dict.	*	*	—	*	*	*
Industrie	JOLY (CILF)	*	*	—	—	—	—
Juridique	Navarre	*	*	—	*	—	—
Management de projet	AFNOR	*	*	—	—	—	—
Nucléaire	FRAMATOME	*	*	—	—	—	—
Routes (AIPCR)	Maison du Dict.	*	*	—	—	—	—
Sciences Harrap		*	*	—	—	—	—
Slang Harrap		*	*	—	—	—	—
Traitement de données	Oldenbourg	*	*	*	—	—	—
Thésaurus :							
SPINES	UNESCO	*	*	—	*	—	*
ILO	BIT	*	*	*	*	—	—
ET BIENTÔT :							
Biologie moléculaire		*	*	—	—	—	—
Chantiers		*	*	—	—	—	—
Chimie		*	*	*	—	—	—
Domotique		*	*	—	—	—	—
Immobilier		*	*	—	—	—	—
Industries graphiques		*	*	—	—	—	—
Machines agricoles		*	*	*	*	*	*
Marine		*	*	—	—	—	—
Mécanique		*	*	*	*	—	—
Police		*	*	—	—	—	—
Sigles		*	*	—	—	—	—
Thésaurus Éducation	BEI	*	*	—	—	—	—
Travaux publics		*	*	—	—	—	—