

## **Patrons de phrase, raccourcis pour apprendre rapidement à parler une nouvelle langue**

Michael Zock, Guy Lapalme

(1) CNRS – LIF (Aix-Marseille II)  
Laboratoire d'Informatique Fondamentale  
Case 901, 163 avenue de Luminy,  
F-13288 Marseille Cedex 9

(2) RALI-DIRO  
Université de Montréal  
CP 6128, Succ. Centre-Ville  
Montréal, QC Canada H3C 3J7  
[michael.zock@lif.univ-mrs.fr](mailto:michael.zock@lif.univ-mrs.fr), [lapalme@iro.umontreal.ca](mailto:lapalme@iro.umontreal.ca)

### **Résumé**

Nous décrivons la création d'un environnement web pour aider des apprenants (adolescents ou adultes) à acquérir les automatismes nécessaires pour produire à un débit "normal" les structures fondamentales d'une langue. Notre point de départ est une base de données de phrases, glanées sur le web ou issues de livres scolaires ou de livres de phrases. Ces phrases ont été généralisées (remplacement de mots par des variables) et indexées en termes de buts pour former une arborescence de patrons. Ces deux astuces permettent de motiver l'usage des patrons et de créer des phrases structurellement identiques à celles rencontrées, tout en étant sémantiquement différentes. Si les notions de 'patrons' ou de 'phrases à trou implicitement typées' ne sont pas nouvelles, le fait de les avoir portées sur ordinateur pour apprendre des langues l'est. Le système étant conçu pour être ouvert, il permet aux utilisateurs, concepteurs ou apprenants, des changements sur de nombreux points importants : le nom des variables, leurs valeurs, le laps de temps entre une question et sa réponse, etc. La version initiale a été développée pour l'anglais et le japonais. Pour tester la généralité de notre approche nous y avons ajouté relativement facilement le français et le chinois.

### **Abstract**

We describe a web application to assist language learners (teenagers or adults) to acquire the needed skills to produce at a 'normal' rate the fundamental structures of a new language, the scope being the survival level. The starting point is a database of sentences gleaned in textbooks, phrasebooks, or the web. We propose to extend the applicability of these structures by generalizing them: concrete sentences becoming productive sentence patterns. In order to produce such generic structures (schemata), we index the sentences in terms of goals, replacing specific elements (words) of the chain by more general terms (variables). This allows the user not only to acquire these structures, but also to express his/her own thoughts. Starting from a communicative goal, he instantiates the variables of the associated schema with words of his choice. We have developed a prototype for English and Japanese, adding Chinese and French without too many problems.

**Mots-clés :** apprentissage de langues, production de langage, livres de phrases, patrons, schéma de phrase, structures fondamentales

**Keywords:** foreign language learning, language production, phrasebook, sentence patterns, basic structure

## 1 Introduction

Produire du langage consiste, schématiquement parlant, à faire dans l'ordre les trois choses suivantes : concevoir un message, le traduire en langue, communiquer ce résultat sous forme graphique ou orale. Ceci semble simple, car tout le monde parle au moins une langue, et deux tiers de la population sur cette planète est bilingue. Pourtant, il n'y a pas l'ombre d'un doute : s'exprimer spontanément et à un débit normal en langue est une tâche difficile. Étant donné une intention de communication ('inviter quelqu'un au restaurant' ; 'raconter un film'), on doit concevoir un, voire plusieurs messages, (conceptualisation), trouver des mots convenables (lexicalisation) insérer ces éléments au bon endroit d'un schéma de phrase à déterminer également (syntaxe), effectuer des flexions et accords (morphologie), prononcer ce résultat (articulation) tout en commençant à planifier le segment suivant (idée).

La tâche n'est donc pas aussi aisée que cela en avait l'air et la difficulté tient essentiellement à trois facteurs : la limitation des ressources du système de traitement (cerveau, surcharge cognitive), la complexité du processus (parallélisme, multitâches) et le volume de données hétérogènes à unifier. En effet, la tâche exige un très grand nombre de choix en très peu de temps. Les informations à traiter sont distribuées à travers plusieurs niveaux. Elles sont donc de nature différente (conceptuel, linguistique, moteur). Les choix peuvent avoir des conséquences multiples, imprévisibles et interdépendantes. Enfin, les éléments à utiliser (faits, mots) doivent être localisés dans un énorme réservoir : base de faits/connaissances (encyclopédie), dictionnaire mental. Si jamais une de ces étapes tarde, ou pire, si la recherche échoue, on assiste à des pauses plus ou moins prononcées, pouvant aller jusqu'au silence total. Ceci peut facilement arriver dans le cas du mot sur le bout de la langue. Imaginez un instant comment trouver un mot particulier parmi les, disons, 30 à 60 000 mots stockés (les chiffres avancés varient selon les auteurs). C'est chercher la fameuse aiguille dans une meule de foin. La performance est impressionnante, équivalant à la consultation manuelle d'un dictionnaire comme *Le Grand Robert* trois fois par seconde pendant plusieurs heures.<sup>1</sup>

Si parler est déjà difficile en langue maternelle, s'exprimer couramment en langue étrangère est une véritable prouesse. Bien entendu, ce n'est pas quelque chose d'inné. Cela a été appris. La question est de savoir comment aider quelqu'un ayant cet objectif. Voici le but de notre travail.

Concernant la méthode, nous poursuivons actuellement deux directions. D'un côté, nous sommes en train de construire une grande *base de phrases multilingues* — environ 40.000 phrases, soit en anglais-japonais (A-J), soit français-japonais (F-J)— de l'autre, nous construisons une *base de phrases d'exercices*, destinée à des apprenants cherchant à acquérir l'habileté nécessaire pour s'exprimer à un débit 'normal' dans une nouvelle langue. Si les phrases de la première base sont assez variées (leur lien étant essentiellement du type 'thématique'), celles de la seconde varient sur très peu d'éléments, ce qui est normal, dans la mesure où leur fonction est de montrer un invariant ou une régularité de la langue.

Les phrases de ces deux bases viennent pour la plupart de livres scolaires, de sites comme Tatoeba (<http://tatoeba.org/fre>) et de *phrasebooks*<sup>2</sup>. Comme chaque couple de langue contient des phrases différentes (les phrases du couple A-J sont différentes de celles du couple F-J), il faut les traduire pour les autres langues, pour permettre ensuite l'accès à partir de n'importe laquelle de ces langues. C'est ce que nous avons commencé à faire, en ajoutant une 4ème langue (Chinois). La traduction faite, on pourra donc non seulement travailler sur chacune des langues en faisant des exercices (travail décrit plus bas), mais également voir la traduction des phrases dans n'importe laquelle de ces langues.

A l'avenir nous aimerions étendre ces deux bases en les enrichissant (plus ou moins) automatiquement, puis établir un pont entre les deux, pour que l'ensemble des phrases puisse être utilisé par l'apprenant pour s'exercer soit en mode traduction, soit en mode production de phrase. Nous présentons ici un générateur d'exercices dont la vocation est d'aider des adolescents ou des adultes à apprendre à s'exprimer à un débit 'normal' en langue étrangère. Le niveau visé étant celui de la survie, nous envisageons un vocabulaire et

<sup>1</sup> Si jamais le chiffre avancé vous paraît élevé, il est bon de savoir que le "Lexique anglais/français des sports olympiques", destiné aux journalistes ayant couvert les jeux de Sidney (2000), contenait déjà pratiquement 14 000 mots, avec 1000 entrées rien que pour les sports aquatiques (rubrique natation).

<sup>2</sup> Un 'phrasebook' est une collection de phrases traduites et organisées par thèmes. Ce type de recueil existe depuis fort longtemps en version papier, et plus récemment sous forme électronique (Fafiotte et al. 2009, Boitet et al. 2007). Ceux intéressés par une version commerciale consulteront <http://speak.econtrader.com/>

des structures en nombre et complexité limités. Avant de décrire notre proposition, voyons deux modèles très influents, caractérisant la production de langage.

## 2 Arrière plan théorique et motivation de notre approche

Comme nous intéressons à la production de langage, ou plus précisément à l'acquisition de cette habileté par des adultes, nous allons présenter les deux principales approches pour les comparer avec la nôtre. Malgré les nombreuses propositions on distingue deux grandes approches : celles proposés par des psychologues (Garret, 1980; Levelt, 1989, Bock, 1995, Ferrand, 2002)<sup>3</sup>, limités généralement à la phrase, et celles venant de la part de linguistes informaticiens (Reiter & Dale, 2000), visant généralement le texte. Bien entendu, les objectifs de ces deux communautés sont assez différents. Les uns s'intéressent au traitement par le cerveau (production de phrases en temps réel)<sup>4</sup>, et les autres s'intéressent au traitement par la machine (TAL). Si les psychologues visent des compromis (traitements imparfaits aux différents niveaux) et la souplesse, les informaticiens visent l'économie et la perfection (production sans fautes).

### 2.1 Le modèle de Garrett

La proposition de Garrett est à la base de pratiquement tous les modèles de production utilisés en psychologie, y compris celui de Levelt (1989). Il consiste principalement en un *conceptualiseur* (message), un *formulateur* (structure linguistique) et un *synthétiseur* de la parole (articulation). A noter qu'on ne passe pas directement du message aux sons, on passe par un module intermédiaire, le composant linguistique, qui joue un rôle de médiateur. C'est d'ailleurs surtout ce module qui a retenu l'attention de Garrett, car les traitements linguistiques laissent des traces. Garrett s'est donc appuyé sur une grande base de données d'erreurs pour construire son modèle.

La tâche du **conceptualiseur** consiste à élaborer un *message* (conceptualisation) afin de réaliser un but ou une intention de communication. Cette structure ou forme de représentation est plus ou moins élaborée, et elle est uniquement conceptuelle. C'est sur elle que s'effectueront les opérations linguistiques qui préciseront alors progressivement cette structure sous-spécifiée. Il y a des bonnes raisons de croire que cette structure est largement sous-spécifiée (Zock 1996). Des contraintes d'espace (mémoire de travail) et de temps (pression de production, manque de temps) sont des facteurs suffisamment contraignants pour dissuader le producteur d'encoder trop en détail le message, car s'il prend trop tôt des engagements forts il s'enferme, réduisant considérablement les options précieuses, utilisables ultérieurement. D'ailleurs, une des astuces rendant possible la production en temps réel est justement de partir d'une structure plus ou moins vide, coquille qu'on enrichira ensuite progressivement.

Le **formulateur** prend en charge des aspects fonctionnels, positionnels et phonologiques des éléments utilisés pour communiquer le message. Le niveau fonctionnel est responsable de l'*encodage grammatical* : les concepts seront remplacés par des mots, ou plus précisément par des lemmes, auxquels on assigne le rôle qu'ils doivent jouer au sein de la phrase. Ainsi faisant on produit une *représentation fonctionnelle* de la phrase. A l'étape suivante (*encodage phonologique*) on détermine la *représentation positionnelle*, c'est-à-dire, on récupère la forme phonologique, les caractéristiques segmentales et prosodiques des lemmes (qui, du coup deviennent des lexèmes) et on spécifie l'ordre des mots en les intégrant dans la structure spécifiée à l'étape précédente. L'**articulateur** doit transformer les symboles du module précédent en sons, afin d'évoquer chez l'auditeur des idées correspondantes à celles ayant donné naissance aux paroles du locuteur.

### 2.2 Le modèle de Reiter et Dale

Le modèle de Reiter et Dale (2000) se décompose également en trois étapes : planification globale, planification locale et formulation. Bien qu'existant dans certains systèmes, la synthèse de la parole, dernier élément de la chaîne, est rarement implémentée.

<sup>3</sup> Pour voir une comparaison des différentes approches on consultera Fromkin (1993), pour des propositions pour un modèle de production bilingue, voir (de Bot, 2000 ; Marini et Fabbro, 2007).

<sup>4</sup> Ce qui implique tout ce qu'on sait des imperfections liées à la performance : surcharge de la mémoire de travail, incapacités d'accéder à une information (latences, 'mot sur le bout de la langue'), interférences (erreurs), incohérences discursives, etc.

La **planification globale** (macroplanning) s'occupe du *choix de contenus* et de la *structuration du document*. Le premier décide des informations à communiquer explicitement dans le texte, en tenant compte des objectifs, connaissances, intérêts et croyances de l'interlocuteur. Le second traite le *groupement* (clustering) et l'*ordonnement* des messages pour produire un ensemble cohérent, tout en évitant des déductions malheureuses (« Elle est tombée enceinte, ils se sont mariés. » vs. « Ils se sont mariés. Elle est tombée enceinte. »). L'ajout de connecteurs peut s'avérer nécessaire afin de révéler le rôle rhétorique des différents fragments (cause vs. concession) : « Il est arrivé juste à temps (car / en dépit) il y avait énormément de circulation. » Les techniques les plus utilisées pour choisir les contenus et déterminer leur organisation sont les 'schémas' (McKeown, 1980),— qui, bien que formellement différentes, sont fonctionnellement équivalentes aux nôtres,— et la 'RST' (Mann/Thomson, 1988).

La **planification locale** (microplanning) comporte la production d'*expressions référentielles* (pronoms), le *choix de mots* (lexicalisation) et l'*agrégation* (harmonisation). La génération d'expressions référentielles consiste à nommer ou à décrire l'objet visé de manière à permettre sa discrimination parmi un ensemble d'alternatives (la voiture, la voiture jaune, celle-là). La lexicalisation consiste à remplacer des concepts par des mots (CHIEN : canin, chiot), et enfin l'agrégation consiste à couper l'espace sémantique (réseau sémantique, représentant l'ensemble des messages à transmettre) pour permettre l'intégration des fragments conceptuels dans le cadre d'un paragraphe ou d'une phrase sans produire une structure trop déséquilibrée. Cette étape peut impliquer l'élimination d'éléments redondants. Les deux ressources les plus importantes à ce stade sont le dictionnaire et la grammaire, la première pour convertir les concepts en mots, et la seconde pour unifier les fragments en phrases.

La **formulation** consiste à convertir des représentations abstraites de phrases en texte concret, à la fois au niveau linguistique (réalisation linguistique) et au niveau de la mise en page (structure de réalisation): des fragments abstraits de texte (sections, paragraphes) sont signalés par des symboles de balisage.

### 2.3 Notre méthode, une approche hybride : les patrons ou schémas de phrases

Comme nous l'avons montré, produire du langage en temps réel est une tâche hautement complexe. Nous présentons ci-dessous une approche, montrant comment l'acquisition d'une telle performance peut néanmoins être rendue possible. Pour voir comment elle se situe par rapport aux travaux mentionnés ci-dessus, nous avons essayé de l'intégrer dans le cadre de Reiter & Dale.

Avant de présenter notre approche, nous soulignons qu'il est hautement improbable que les locuteurs ou apprenants passent par toutes les étapes décrites, appliquant une à une les règles ou contraintes décrites par des linguistes dans leurs grammaires formelles. Il y a au moins trois raisons qui nous poussent à douter de cela :

- raison liée à la *mémoire* : les gens n'ont pas stocké dans leur mémoire l'ensemble des règles décrites par les grammairiens. Essayez donc d'évoquer une des ces règles hormis celles concernant les accords. D'ailleurs, même si on avait stocké ces règles, on ne pourrait pas les utiliser séquentiellement, car leur ordre variera en fonction des informations conceptuelles qui nous viennent à l'esprit dans un ordre quelconque. La mémoire de travail (Baddeley, 1992) est déjà très sollicitée par d'autres tâches, notamment, l'encodage du message;
- raison d'*attention* : on ne peut se concentrer que sur un petit nombre de tâches (ou d'objets) à la fois. Les capacités de traitements parallèles sont sûrement bien moindres en cas d'apprentissage d'une nouvelle langue, comparées à la langue maternelle pour laquelle les mécanismes sont déjà bien rodés.
- raison de *temps* : la conception de message et sa traduction en langue, sont des processus extrêmement rapides. Tous les locuteurs le savent bien, une idée non exprimée à temps risque de retomber dans l'oubli, d'où une course effrénée entre les idées et leur expression. Les locuteurs doivent donc rapidement traduire leurs messages ; chercher les règles à appliquer et les appliquer serait beaucoup trop long.

Les linguistes décrivent généralement les langues en termes de règles, mais la plupart des gens n'apprennent pour ainsi dire jamais de telles descriptions. Il est encore moins probable qu'ils les appliquent toutes, encore moins pendant les phases initiales de l'apprentissage d'une nouvelle langue. En revanche, les gens apprennent des modèles conformes à des règles. Et s'ils utilisent des règles, c'est essentiellement

localement (morphologie), sachant que le gros du travail a été fait au préalable par le choix de schémas de phrases. Ces derniers peuvent d'ailleurs être vus comme une prise de vue instantanée d'un processus dérivationnel.

Les modèles sont des invariants, c'est-à-dire des abstractions faites à partir des formes que sont les phrases concrètes. Leurs composants (mots) peuvent être caractérisés en divers termes (syntaxiques, sémantiques, les deux). Autrement dit, il y a plusieurs manières de caractériser la même chaîne. Il n'y a pas de caractérisation absolue. Tout dépend du point de vue et du niveau d'abstraction. Les modèles ou schémas de phrase sont des structures, susceptibles d'être construites dynamiquement via des règles. Cependant, on peut également les concevoir comme des unités, structures holistiques, rencontrées, stockées ou récupérées telles quelles. Un locuteur performant posséderait donc une grande librairie de modèles et un bon index lui permettant de localiser rapidement le schéma de phrase nécessaire. Au fond, c'est un peu comme s'il cherchait un mot dans un dictionnaire ou un thésaurus. Autrement dit, les dictionnaires de mots et les mémoires de (schémas de) phrases se ressemblent. Ce sont des bases de données indexées, espaces balisés, dans lesquels on navigue pour récupérer l'élément nécessaire.

Récupérer d'un seul coup toute une phrase (ou presque) nous épargne des efforts de calcul tout en nous faisant gagner du temps. C'est d'ailleurs probablement la raison pour laquelle tant de gens s'en servent en langue (apprenants, interprètes, journalistes, etc.) ou dans d'autres domaines (musique, programmation, jeu d'échecs, etc.). Bien sûr, il y a un prix à payer : les modèles doivent non seulement être accessibles (voir ci-dessous), ils doivent également être adéquats et compatibles avec l'idée à exprimer. Si les modèles ont des qualités, ils ont aussi des faiblesses : ils sont rigides, et ils occupent de la place mémoire. Il faut donc procéder parcimonieusement. Toute variation linguistique ne justifie pas forcément qu'on en fasse une abstraction. Imaginez les variations morphologiques (temps). Elles ne méritent guère qu'on en tienne compte dans des schémas de phrase. Prenez, par exemple, les deux phrases suivantes et leurs schémas respectifs : (A) Je vais à New York cet été [je vais < LIEU > < TEMPS >]. --- (B) J'ai été à Madrid la semaine dernière [j'ai été < LIEU > < TEMPS >].

Vue la similitude des deux phrases il n'est guère justifiable d'abstraire deux schémas différents. Il serait beaucoup plus raisonnable d'avoir un modèle rendant compte de la *structure globale* [PERSONNE aller < LIEU > < TEMPS >] et un ensemble de paramètres (règles) prenant en charge des ajustements locaux : accords (singulier/pluriel), flexion de verbes (passé, présent), etc.

Tout comme les schémas, les règles ont certains inconvénients. Certes, elles ont la puissance nécessaire pour rendre compte de l'expressivité de la langue (ensemble de variations légalement possibles), mais, vu leur granularité, leur nombre devient prohibitif, empêchant le locuteur de faire son travail à temps. C'est pourquoi nous suggérons une approche hybride, approche à deux vitesses : des *schémas* pour les structures globales et des *règles* pour les ajustements morphologiques (niveau local). Cette combinaison offre le meilleur compromis. D'une part elle permet de minimiser les besoins de calcul (allégeant du coup la mémoire et l'attention), d'autre part elle maximise la puissance (rapidité pour faire le gros du travail) et la flexibilité. Cette dernière est requise pour effectuer des ajustements (accords) ou des restructurations locales.

Lorsqu'on apprend une nouvelle langue, on apprend généralement une liste de mots (vocabulaire) et un mécanisme de construction de phrase (grammaire). Spécifiant les combinaisons légales, la grammaire fournit les structures possibles, dans lesquelles le locuteur va insérer les mots choisis pour exprimer (une partie de) ses idées (concepts). Ceci dit, la structure syntaxique peut être obtenue de différentes manières : par le biais d'une construction incrémentale (unification progressive des éléments) ou par une recherche dans la mémoire, auquel cas on la récupère en un seul bloc (schémas).

Bien qu'il s'agisse d'un continuum, on peut imaginer trois grandes approches : (a) l'unité du traitement est le concept ou sa traduction, le *mot* ; (b) l'unité est un *segment de phrase* (phrases lexicales)<sup>5</sup> ou (c) l'unité est toute la *phrase*. La première solution est la plus risquée et la plus coûteuse, car produire des phrases à partir d'unités aussi petites implique une vision très réduite et de nombreux calculs (opérations d'unification). La dernière approche est la plus rapide, souvent la plus sûre, mais, à terme, aussi la plus limitée. Réutilisant une forme telle quelle (imitation d'une phrase), et n'ayant fait aucune abstraction, on

<sup>5</sup> De tels segments sont assez évidents dans des formules ('Veuillez agréer, cher Monsieur, l'expression de mes sentiments distingués'; 'je vous en prie') ou dans des expressions comme : 'dans la mesure où', ou 'Qui aurait pu croire que <phrase> ?', etc. (Nattinger et Decarrico, 1992, Becker, 1975).

reste prisonnier des formes rencontrées. On ne peut exprimer rien d'autre. La stratégie (b) est à terme la plus utile. Bien qu'elle ne soit pas parfaite, la spécificité se payant au prix de la généralité, elle offre néanmoins un excellent compromis entre la vitesse, la puissance et la souplesse. En effet, elle permet d'exprimer rapidement des idées très variées, sans obliger le locuteur à effectuer de nombreux calculs. Ayant récupéré des grands blocs ('chunks' conceptuels lexicalisés) il les insère dans un schéma plus large. Cette stratégie est très utilisée par des interprètes de conférence, car ils travaillent constamment à la limite de la surcharge cognitive. Aussi, au lieu d'attendre la fin de la phrase, ils ont tendance à commencer la traduction le plus tôt possible, opérant sur des fragments plutôt que sur l'ensemble des éléments de la phrase. Cela permet de minimiser la charge mémorielle tout en augmentant le temps disponible pour la partie à venir, partie encore à traduire. Un interprète essaie donc à tout prix de se 'débarrasser' le plus vite possible d'une partie du message, pour avoir le maximum de temps pour la partie restante. D'ailleurs, les interprètes craignent généralement moins les locuteurs à grand débit que ceux au débit lent, ou ceux utilisant des structures emboîtées, car dans les deux cas on a du mal à faire de bonnes prédictions concernant le rôle joué par certains éléments à traduire.

Ceci dit, si la stratégie basée sur les *phrases lexicales* est séduisante, elle est trop ambitieuse, parce que trop difficile, pour un débutant. Ce dernier doit avoir rapidement du succès pour devenir confiant. C'est pour cette raison que nous proposons de travailler avec des phrases plus ou moins toutes faites. Certes, il ne s'agit pas d'apprendre par cœur ces phrases, mais plutôt d'abstraire le schéma sous-jacent, pour pouvoir produire des phrases analogues ou similaires.<sup>7</sup> Autrement dit, nous visons la productivité de la langue. Mais pour y arriver nous essayons de réduire la charge cognitive, tout en visant l'augmentation du contrôle : on ne se concentre que sur un élément à la fois.

Pour voir comment notre modèle se situe par rapport aux autres modèles de production, nous présentons les différentes étapes en termes de ces architectures. Ayant indexé les structures à apprendre en termes de but, l'utilisateur part de celui-ci pour communiquer son intention de communication (1°). Le système répond avec un, voire plusieurs schémas de phrase, parmi lesquels l'utilisateur choisit (2°). Cette étape correspond au *niveau macro* du modèle de Reiter et Dale. En effectuant des choix lexicaux (3°) on est désormais au *niveau micro*, nommé *formulateur* dans le modèle de Garrett. On notera que le choix lexical peut être fait par le système, et c'est comme ça que les choses se passent dans les fameux 'pattern drills' (Besse, 1975, Le Rouzo, 1975).

NIVEAUX	ENTRÉES CONCEPTUELLES	SORTIES LINGUISTIQUES
NIVEAU GLOBAL (macro)	1° <i>but</i> de communication :  comparaison	Ensemble de schémas de phrase  (a) < OBJET <sub>1</sub> > être plus < ATTRIBUT > que < OBJET <sub>2</sub> >  (b) < OBJET <sub>2</sub> > être moins < ATTRIBUT > que < OBJET <sub>1</sub> >
	2° <i>cadre syntaxique</i> : (a)	< OBJET <sub>1</sub> > être plus < ATTRIBUT > que < OBJET <sub>2</sub> >
NIVEAU LOCAL (micro)	3° valeurs <i>lexicales</i> :  OBJET <sub>1</sub> = eau ; OBJET <sub>2</sub> = vin ATTRIBUT = cher	Structure lexicalement spécifiée :  eau être plus cher que vin.
	4° valeurs <i>morphologiques</i> :  OBJET <sub>1</sub> = singulier OBJET <sub>2</sub> = singulier	Structure conceptuelle et linguistique complète  L'eau est plus chère que le vin

Tableau 1 : L'entrée conceptuelle, un processus en quatre étapes

Ceci dit, si l'on souhaite un système ouvert, donc sensible aux besoins de l'utilisateur, on laisse le choix à ce dernier. La structure choisie en (2°) sera instanciée par la valeur lexicale (3°), puis (4°) complétée par des valeurs morphologiques (nombre, temps, etc.). Désormais le système a tout ce dont il a besoin pour produire la phrase. On notera que l'entrée conceptuelle est répartie sur trois niveaux : au niveau le plus

<sup>7</sup> L'analogie est un principe d'apprentissage bien connu, utilisé par des enfants pour apprendre les régularités d'une langue (Berko, 1958). Elle a été proposée comme base dans des 'exercices structuraux' ('pattern drills' ou ' patrons de phrases'), et même en traduction automatique (Nagao, 1984), bien qu'il s'agissait là en fait plutôt de similarités.

profond elle consiste à choisir un schéma global (schéma de phrase) via un but. Aux niveaux suivants, on spécifie respectivement les valeurs lexicales et morphologiques (nombre, temps). Ainsi faisant on affine peu à peu un message sous-spécifié. Ce type de décomposition a plusieurs avantages. Les informations ne sont demandées que lorsqu'elles sont pertinentes ou nécessaires, ce qui réduit la complexité du traitement, tout en augmentant son contrôle. Deuxièmement, cette méthode est bien plus rapide pour transmettre un message que de naviguer dans une immense ontologie lexicale ou conceptuelle, tel que cela a été suggéré ailleurs (Power et al, 1998 ; Zock, 1991; Zock et al 2009).

### 3 Construction de la ressource

Étant donné le processus décrit dans le tableau 1, nous avons séparé la description des *but*s de celle des méthodes permettant de les atteindre, les *patrons* ou structures de phrases. Certains éléments des phrases ont été généralisés via des variables syntaxiques ou lexicales. Par exemple, l'objectif 'se présenter' peut être atteint via une des deux structures suivantes : «On m'appelle X" ou "Mon nom est X" auquel cas la variable X peut être instanciée par le nom, prénom, ou petit nom de l'orateur. Les valeurs autorisées peuvent être définies dans un dictionnaire. Pour un même objectif, on peut imaginer un grand nombre de modèles et de types de variables, et ceci pour différentes langues.

Étant donné la nature hiérarchique des objectifs et des patrons de phrase, nous avons décidé d'utiliser une structure arborescente codée en XML pour conserver l'information. Avec un éditeur approprié, un linguiste peut facilement ajouter de nouveaux objectifs, des modèles de phrases associés, des variables et des entrées lexicales sans avoir à s'occuper de la complexité du programme informatique qui affiche le résultat à l'utilisateur. Le programme de traitement lit cette structure et l'interprète à la volée pour générer des patrons et des phrases qui sont ensuite présentées à l'utilisateur et comparées avec ce que celui-ci dit ou écrit.

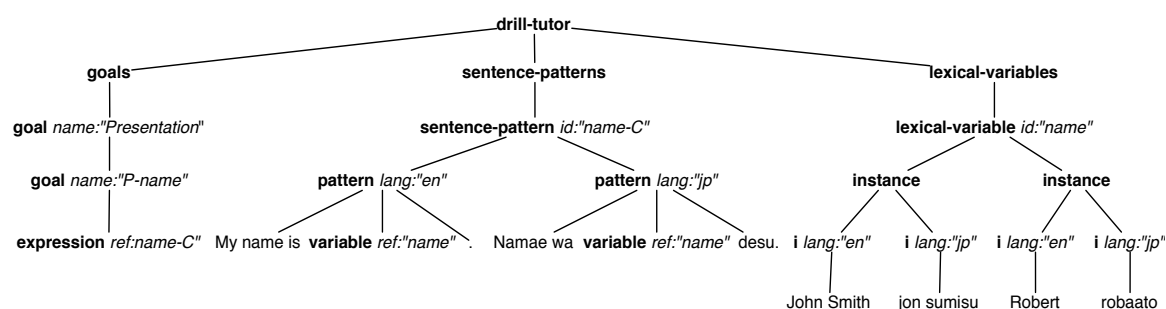


Figure 1: Définitions de *but*s, *schémas de phrases* et *entrées lexicales* sous forme d'arbre.

### 4 Aspects techniques et l'état actuel du prototype

Pour réaliser notre 'entraîneur' de langue (DrillTutor), nous avons décidé d'utiliser un navigateur Web à cause de sa capacité d'affichage de plusieurs jeux de caractères et de sa disponibilité sur toutes les plateformes. Pour créer les documents nous utilisons PHP, car il fournit un moyen assez commode pour créer dynamiquement des pages web et pour accéder aux fichiers XML. Pour afficher la prononciation des mots japonais, nous avons utilisé deux scripts : le romaji (une forme de représentation phonétique proche de l'anglais) et l'hiragana, le syllabaire principal du Japonais. La conversion est faite automatiquement via une fonction PHP qui ensuite affiche le résultat dans le browser.

Une fois l'objectif sélectionné, le programme engendre une phrase en remplaçant dans le modèle les valeurs des variables lexicales. Une variable peut être utilisée plusieurs fois à condition d'avoir été définie. Sa valeur sera alors propagée à travers les différents modèles de phrase. Comme une même valeur lexicale peut correspondre à différentes variables, ce lexème recevra des valeurs distinctes. Cette fonction est utilisée dans les modèles comme : « Êtes-vous Chinois? Non, je ne suis pas Chinois, mais Japonais » où Chinois et Japonais sont considérés comme des variables dans le modèle. Les deux premières occurrences auraient le même nom, alors que la dernière serait nommée différemment.

Côté serveur, on garde dans un fichier des statistiques concernant la performance des utilisateurs. Celle-ci

peut être utilisée pour déterminer la sélection des variables d'un modèle et la manière d'afficher des buts. Ceci dit, à l'heure actuelle, on choisit au hasard les variables des structures de phrases et on affiche tous les buts.

Nous nous servons de Javascript pour obtenir une interaction locale à l'intérieur des pages. Par exemple, comme on voit sur la figure 2b, la page web contient initialement tous les buts et modèles de phrase. Cependant l'affichage est contrôlé localement via Javascript qui lui gère les clics de la souris et le contenu du champ texte pour afficher un menu de mots-clés suggérés pour trouver des buts. L'interface du système (à savoir, les titres et incitations du système) se trouvent dans un autre fichier XML qui stocke les chaînes de message dans différentes langues (actuellement Français et Anglais).

La séparation de la description des objectifs et des patterns permettra à terme l'ajout par l'utilisateur de nouveaux buts et schémas de phrase. Cette fonctionnalité n'ayant pas encore été mise en œuvre, les objectifs et les schémas doivent être saisis à l'aide d'un éditeur de texte XML. Un schéma XML permet de valider que l'information sur les buts, les patrons et les variables est codée dans le bon format et que toutes les références à des variables et les schémas sont bel et bien des éléments présents dans le fichier XML. Initialement, nous avons défini les schémas en trois langues (anglais, français et japonais). Depuis lors des buts, des schémas et des valeurs lexicales ont été ajoutés en chinois, et ceci en très peu de temps. Le système a une quinzaine de buts et de schémas et 150 éléments lexicaux, en quatre langues dont deux langues d'interface. Bien que ces chiffres puissent sembler très petits, il faut garder à l'esprit les points suivants. Un des objectifs de notre travail était de vérifier les difficultés que présenterait l'ajout d'une nouvelle langue. Il s'est avéré que même l'ajout d'une langue typologiquement aussi différente que le chinois, ne présentait pas un obstacle majeur. En fait, le nombre de schémas et la taille du vocabulaire ne sont pas vraiment des facteurs qui comptent actuellement. L'accent a été mis sur la réalisation d'un éditeur conçu pour créer, modifier et utiliser une base de données. Cette dernière peut être facilement étendue. L'architecture que nous avons définie permet de gérer autant de buts, de schémas et d'éléments lexicaux que l'on veut. Il suffit pour cela d'éditer des fichiers XML en suivant un schéma bien défini (structure), et en veillant à la bonne formation, ce qui peut être fait via un éditeur de schéma XML.

Nous prévoyons de développer les composants suivants dans notre entraîneur de langue, les parties 2 et 4-8 n'étant pas encore implantées:

1. Librairie de schémas et scénarios indexés en termes de buts ;
2. Dictionnaires multilingues ;
3. Translitérateurs romaji et hiragana ;
4. Composants morphologiques : générateur de formes et d'accord ;
5. Synthèse de la parole ;
6. Vérification de cohérence et de bonne formation ;
7. Induction de schémas via une ressource comme WordNet ;
8. Extraction d'exemples via un corpus ;
9. Paramètres concernant les choix et les performances de l'utilisateur.

## 5 Exemple d'utilisation

Le système est disponible à l'adresse : <http://agil.lif.univ-mrs.fr/DrillTutor>. L'utilisateur choisit le but de la communication. Les objectifs étant indexés et ordonnés de manière hiérarchique, l'utilisateur peut les atteindre soit en écrivant leur nom dans une case réservée à cet effet, soit en naviguant dans l'arborescence (figure 2a). Pour un objectif donné il suffit de cliquer sur son nom pour voir se développer soit les sous-but, soit la structure associée, les variables étant indiquées en gras. On est alors arrivé au niveau des feuilles. Dans la figure 2a, l'utilisateur a demandé le schéma permettant d'atteindre le but "origine". Après avoir choisi dans quelle langue il aimerait faire l'exercice, en l'occurrence le Japonais, il passe par les deux étapes suivantes. D'abord il spécifie la valeur de la variable 'origine' (2b) puis il vérifie si son résultat correspond à celui de la machine (2c).



Drill Tutor Goals and Correspondences

Welcome Bob

Click on a *FR*, *JP* or *CN* link to start exercising a goal.  
Hover the mouse on a bold word to see alternatives

Search for a goal

- Presentation
  - name
  - origin
    - I'm **nationality**. [FR](#) [JP](#) [CN](#)
    - title name origin
- Question Answer
  - Counting
  - name pos
  - origin neg
  - name origin neg
  - time

French, English, German, Australian, Japanese, Indonesian, Filipino, Thai

Figure 2a : **Choix du but** (ici, 'origine'), d'une structure lorsqu'il y en a plusieurs et de la *langue* à apprendre (**FR** : français, **JP** : japonais; **CN** : chinois), ici, japonais

## Your goals : Presentation - origin

Thanks Bob. Now, translate this sentence into Japanese.

I'm French.

Hit return and compare your answer with ours.

Figure 2b : **Entrée**, valeur de la variable *nationalité* = française

Watashi wa furansu jin desu.  
わたしわふらんすじんです。

Did you get it right?  yes  no

Select a different goal

S

Figure 2c : **Résultat** en lettres romanes et en script japonais <sup>8</sup>

<sup>8</sup> A noter, que la *particule* 'wa', tout juste après 'watashi' ('je') indique le rôle (thème) du pronom personnel. Elle devrait être écrite 'は', équivalent de la syllabe 'ho'. Mais comme on prononce les deux syllabes de la même manière ('**watashi wa**'), notre translérateur produit dans les deux cas le même caractère ('わ'). Ceci devrait être corrigé, mais cela demande un programme capable de reconnaître les différentes fonctions d'un mot ou d'une syllabe. A noter également, que les noms propres (noms de personnes et de villes) devraient être écrits en kanji ou en katakana. Comme il s'agit d'acquérir un niveau de base à l'oral, nous avons ignoré ces deux subtilités, car elles rendraient notre tâche et celle des apprenants inutilement difficiles.

## 6 Discussion

S'agissant d'un travail en cours, notre proposition est forcément incomplète. Nous n'avons pas fait référence à une théorie d'apprentissage particulière, car nous ne cherchons pas à enseigner tous les aspects de la langue, ou ceux de son traitement. Nous nous sommes limités à un seul aspect, la production de phrases simples. Nous sommes convaincus que l'exercice est indispensable et qu'il doit avoir certaines caractéristiques — limitation des éléments à faire varier à l'intérieur des structures à apprendre — pour permettre à l'apprenant d'atteindre le niveau suivant, à savoir, la construction spontanée de phrases complexes (voir également Levelt, 1970; Krashen, 1981). Ceci veut dire, mener en parallèle et de manière incrémentale l'encodage du message et son expression (forme linguistique). Malgré le grand nombre de travaux en psycholinguistique, en linguistique informatique et en didactique —(Chapelle, 2001; 2003, Crystal; 2001, Warschauer, 1998, Gethin et Gunnemark, 1995; Holland, Kaplan et Sam, 1995, Swartz et Yazdani, 1991, Brown, 1987, Novak et Gowin, 1984, Wilkins, 1972;)— il n'y a pratiquement rien qui permette leur transposition pour apprendre à passer d'un *but* et d'un *message* vers son *expression*. Ceci vaut à la fois pour les prototypes issus des laboratoires de recherche et pour les produits industriels (Pimsleur, TellMeMore, Rosetta Stone, etc.).

Il est intéressant de noter que l'utilisateur 'apprend' pratiquement les mêmes connaissances qui nous ont guidés à construire notre générateur. Celui-ci est simple, voire rudimentaire, mais assez facile à mettre en oeuvre et à s'approprier ou apprendre. En suivant la démarche proposée par notre programme on apprend donc non seulement à engendrer des phrases, mais aussi comment procéder : la nature et l'ordre des informations à fournir et à traiter sont suggérés par le programme. Bien entendu, on peut imaginer d'autres situations et d'autres stratégies. Nous considérons justement la nôtre comme une étape préliminaire, préparant l'apprenant à celle d'une génération plus souple, mais demandant plus de connaissances. C'est elle qu'on utilise normalement en discours spontané (imaginez que vous deviez décrire un film que vous venez de voir, tâche très exigeante, notamment au niveau conceptuel). D'ailleurs, rien n'interdit de commencer par le choix des mots avant de trouver leur place dans un cadre syntaxique. Mais ceci suppose quelque chose ressemblant à une grammaire d'unification en arrière-plan, type de connaissance qu'on n'acquiert que plus tard et progressivement, lorsqu'on a rencontré un très grand nombre de phrases.

Nous sommes limités à un accès simple aux phrases à partir des buts. Or toutes les phrases ne peuvent pas être indexées, et bon nombre d'entre elles peuvent être indexées selon différents points de vue. Ceci dit, étant donné notre objectif (niveau de survie), on peut raisonnablement supposer que les phrases puissent jouer le rôle que nous leur avons assigné.

Plusieurs problèmes d'organisation des schémas se posent, à savoir, quels éléments de la phrase généraliser et en quels termes. Quant au nom de la variable utilisée pour remplacer un terme (mot, expression), c'est à la fois un problème de métalangage (est-ce que 'mouton', issu de 'dessine-moi un mouton', doit être remplacé par 'animal', 'groupe nominal' ou par quel autre terme?) et un problème d'inclusion. Plus le terme choisi est large, plus le nombre de valeurs possibles est grand ('voici un <objet>' vs 'voici un <fruit>'). Nous avons induit les schémas manuellement, mais on pourrait imaginer d'automatiser l'induction de schémas, en remplaçant les mots par leur hyperonyme via une ressource comme WordNet.

Un des aspects le plus prometteurs, mais aussi des plus difficiles, est la perspective de créer ou d'étendre (semi-)automatiquement la base de phrases et celle de schémas. On allégerait ainsi le travail du créateur, on augmenterait l'étendue ou l'expérience linguistique de l'apprenant tant par la variété pour combattre la monotonie que l'authenticité des exemples. Le défi étant de trouver des exemples qui répondent aux besoins pédagogiques et cognitifs de l'apprenant.

Notre prototype ne comporte pour l'instant ni dictionnaire (mono- ou multilingue), ni générateur de morphologie. Ayant été motivés par la construction rapide d'un prototype, nous avons négligé cet aspect comme celui de l'évaluation qui reste nécessaire, mais reste tributaire d'une implantation plus complète que ne l'est notre prototype actuel.

## 7 Conclusion

Nous avons présenté une base de phrases multilingue convertie en générateur d'exercices pour aider les utilisateurs à acquérir une certaine maîtrise orale en langue étrangère. L'apprentissage des mots et de leurs combinaisons est indispensable pour s'exprimer à un débit 'normal' dans une nouvelle langue. L'acquisition de ces automatismes nous paraît capitale pour accéder à l'étape suivante : savoir produire spontanément des phrases plus élaborées. Vu la vitesse avec laquelle il faut effectuer un très grand nombre d'opérations, il est souhaitable d'automatiser celles qui sont les plus stables, à savoir, les opérations linguistiques (syntaxe, morphologie), car le sens peut varier à l'infini.

Notre approche peut être caractérisée par son ouverture, sa généralité et son extensibilité : de nouvelles informations peuvent être ajoutées à tout moment, et d'autres langues peuvent être mises en œuvre très rapidement. En effet, nous avons testé l'extensibilité de l'outil, constatant qu'il était assez facile d'ajouter de nouvelles informations, ou même une langue typologiquement très différente. Considérant que la production de phrases seulement via des règles était trop lourde, et que les modèles seuls étaient trop rigides et trop gourmands en mémoire (stockage, accès), nous avons opté pour un compromis (processus à deux vitesses) : utiliser des patrons au niveau global et des règles pour les ajustements locaux. Ceci augmente considérablement la vitesse de production de phrases, tout en minimisant le besoin de stockage (patrons, règles).

Bien que la couverture actuelle soit encore très faible et bien que le système n'a pas été évalué, nous pensons que l'approche est viable. Bien entendu, le juge ultime sera l'utilisateur. Va-t-il utiliser un tel système ? Ce dernier lui permettra-t-il d'atteindre les objectifs fixés ? Voici des questions auxquelles nous devons répondre.

## Références

- BADDELEY, A. (1992). Working memory. *Science*, 255, 556-559
- BECKER, J. (1975). The Phrasal Lexicon. In *Proceedings Interdisciplinary Workshop on Theoretical Issues in Natural Language Processing*. Cambridge, Massachusetts June 70-73.
- BERKO, J. (1958). The Child's Learning of English Morphology. *Word*, 14, 150-177
- BESSE, H. (1975) De la pratique aux théories des exercices structuraux. *Etudes de Linguistique Appliquée*, 20, 8-30. Paris, Didier
- BOCK, J.K. (1995). *Sentence production: From mind to mouth*. In J. L. Miller & P.D. Eimas (Ed.), *Handbook of perception and cognition*. Vol. 11: Speech, language and communication. Orlando, FL: Academic Press.
- BOITET, C., BHATTACHARYYA, P., BLANC, E., MEENA, BOUDDH, S., FAFIOTTE, G. & FALAISE, A. & V. VACCHANI (2007). Building Hindi-French-English-UNL resources for SurvTra-CIFLI, a linguistic survival system under construction, actes de *SNLP*, Pattaya, Thaïlande
- BROWN, H. 1987 *Principles of Language Learning and Language Teaching*. Englewood Cliffs, NJ: Prentice-Hall.
- CHAPELLE, C. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing and research*. CUP.
- CRYSTAL, D. (2001). *Language and the Internet*. CUP.
- de BOT, K. (2000). A bilingual production model: Levelt's 'speaking' model adapted. In L. Wei (Ed.), *The bilingualism reader* (pp. 420-442). London; New York: Routledge. (Reprinted from *Applied Linguistics*, 13, 1992, 1-24.)
- FAFIOTTE, G. FALAISE, A. & J. GOULIAN. (2009). CIFLI-SurvTra, deux facettes : démonstrateur de composants de TA fondée sur UNL, et phrasebook multilingue, actes de *TALN*, Senlis
- FERRAND, L. (2002). Les modèles de la production de la parole. In M. Fayol (Ed.), *Production du langage*. *Traité des Sciences Cognitives* (pp. 27-44). Paris: Hermès.

- FROMKIN, V. (1993). *Speech Production*. In Psycholinguistics. J. Berko Gleason & N. Bernstein Ratner, Eds. Fort Worth, TX: Harcourt, Brace, Jovanovich.
- GARRETT, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (pp. 177-220). London: Academic Press
- GETHIN, A. & E. GUNNEMARK. 1995. *The Art and Science of Learning Languages*. Intellect books, Headington, Oxford.
- HOLLAND, M., J. KAPLAN & M. SAMS (Eds.).1995.*Intelligent Language Tutors*. Hillsdale, NJ.: Lawrence Erlbaum Associates.
- KRASHEN, S. (1981). [http://www.sdkrashen.com/SL\\_Acquisition\\_and\\_Learning/index.html](http://www.sdkrashen.com/SL_Acquisition_and_Learning/index.html), *Second Language Acquisition and Second Language Learning*.
- LE ROUZO, M. L. (1975). Y a-t-il une justification psychologique à la pratique des exercices structuraux ? *Etudes de Linguistique Appliquée*, 20, 37-51. Paris, Didier
- LEVELT, W. (1970). Skill theory and language teaching. *Studies in Second Language Acquisition*, 1(1), 53–70.
- LEVELT, W. (1989). *Speaking*. MIT Press, Cambridge, Mass.
- Lexique anglais/français des sports olympiques: jeux d'été, *Insep Publications*, Paris (2000).
- MANN, W. C. & THOMPSON, S. A. (1988). Rhetorical Structure Theory: Toward a Functional Theory of Text Organization', *Text* 8(3), 243--281.
- MARINI, A. and FABBRO, F. (2007) "Psycholinguistic models of speech production in Bilingualism and Multilingualism". In Ardila, A. and Ramos, E. (eds.) "Speech and language disorders in Bilinguals". Nova Science Publishers Inc. New York, NY, pp. 47-67s
- MCKEOWN, K. (1985). Discourse Strategies for Generating Natural-Language Text. *Artificial Intelligence*, 27, 1-41
- NAGAO, M. (1984). A framework of a mechanical translation between Japanese and English by analogy principle. In A. Elithorn & R. Banerji (Eds.), *Artificial and Human Intelligence* (pp. 173–180). Amsterdam: Elsevier.
- NATTINGER, J. et S. DECARRICO, J. (1992) *Lexical Phrases and Language Teaching*, Oxford, Oxford University Press
- NOVAK, J. & D. GOWIN. 1984. *Learning how to Learn*. Cambridge. Cambridge University Press.
- POWER, R., SCOTT, D., & EVANS, R. (1998). What you see is what you meant: direct knowledge editings with natural language feedback. In H. Prade (Ed.), *13th European conference on artificial intelligence (ECAI'98)* (pp. 677–681). Chichester: Wiley.
- REITER, E., & DALE, R. (2000). *Building natural language generation systems*. Cambridge: Cambridge University Press.
- SWARTZ, M. et M. YAZDANI (Eds.). (1991). *Intelligent Tutoring Systems for Foreign Language Learning: The Bridge to International Communication*. Springer Verlag, Berlin.
- WARSCHAUER, M. & HEALEY, D. (1998). Computers and Language Learning: An overview. *Language Teaching*, 31, 57–71. <http://www.lll.hawaii.edu/web/faculty/markw/overview.htm>
- WILKINS D. (1972): *Linguistics and Language Teaching*. London: Edward Arnold.
- ZOCK, M. (1991). Swim or sink: the problem of communicating thought. In M. Swartz & M. Yazdani (Eds.), *Intelligent tutoring systems for foreign language learning* (pp. 235–247). New York: Springer.
- ZOCK, M. (1996). The Power of Words in Message Planning, *16th International Conference on Computational Linguistics*, Copenhagen, pp. 990-995
- ZOCK, M., SABATIER, P. and L. JAKUBIEC. Message Composition Based on Concepts and Goals. *International Journal of Speech Technology*, 11(3-4):181–193, 2008.