

## L'intelligence artificielle et l'étude du système linguistique

*Entre contraintes et facilités d'acquisition des connaissances*

### Artificial Intelligence and the Study of Linguistic Systems

*Between Constraints and Opportunities for Knowledge Acquisition*

**Dre Aicha Ibtissem BOUACHA**

Auteur correspondant, Université Alger 2, [bouacha23ibtissem@gmail.com](mailto:bouacha23ibtissem@gmail.com)

**Soumission : 21.04.2026 – Acceptation : 23.04.2026 – Publication : 01.05.2026**

**Résumé** — L'évolution aidant, nous assistons aujourd'hui à l'émergence de l'intelligence artificielle (IA). Ce concept, censé être un processus d'imitation de l'intelligence humaine devrait servir, comme outil de traitement et d'acquisition des connaissances en matière d'enseignement-apprentissage des langues étrangères. Il devrait ainsi représenter en soi un défi sans précédent pour le système éducatif traditionnel.

Néanmoins, si l'enseignement des sciences dures est de nature conceptuelle et repose sur la création et l'application d'algorithme, *l'enseignement des langues relevant de pratiques sociales est-il en mesure d'accepter ce genre de traitement ?* Si, par ailleurs, difficultés il y a, *quelles sont les contraintes qui empêchent l'application des algorithmes, notamment lors du traitement d'une discipline comme la grammaire ?*

Nous tenterons, dans ce travail, de montrer les contraintes qui entravent l'application des procédés algorithmiques sur la construction d'un savoir linguistique et empêcheraient l'acquisition de connaissances.

**Mots-clés** : *enseignement du FLE, grammaire, intelligence artificielle, approche algorithmique, connaissances.*

**Abstract** — As technology advances, we are now witnessing the emergence of artificial intelligence (AI). This concept, intended to mimic human intelligence, is expected to serve as a tool for processing and acquiring knowledge in the field of foreign language teaching and learning. It is therefore likely to pose an unprecedented challenge to the traditional education system.

Nevertheless, whilst the teaching of hard sciences is conceptual in nature and relies on the creation and application of algorithms, *is the teaching of languages – which falls within the realm of social practices – capable of accommodating this kind of approach?* If there are difficulties, *what are the constraints preventing the application of algorithms, particularly when dealing with a discipline such as grammar?*

In this paper, we shall attempt to highlight the constraints that hinder the application of algorithmic processes to the construction of linguistic knowledge and would prevent the acquisition of such knowledge.

**Keywords:** *Teaching French As A Foreign Language, Grammar, Artificial Intelligence, Algorithmic Approach, Knowledge.*

## Introduction

Depuis les origines de la pensée linguistique, le caractère « scientifique » du concept de *langue* au travers de sa grammaire a toujours été relativisé dans sa conception en tant que concept scientifique abouti : *le concept de science correspond plus à la démonstration mathématique, aux matières de physique, de chimie ou encore au domaine technique qu'aux descriptions empiriques de la langue.*

Ainsi, si l'enseignement des sciences dures telles les mathématiques, les sciences physiques et aussi les matières techniques sont de nature conceptuelle et repose sur la création et l'application d'algorithme, ce qui pourrait s'envisager avec l'utilisation de l'intelligence artificielle puisque celle-ci repose sur des données formelles exécutées dans un environnement informatique dynamique, l'enseignement des langues repose sur des pratiques sociales se traduisant par une diversité linguistique et surtout des situations complexes et trop fluctuantes mettant à rude épreuve toute tentative à vouloir structurer formellement et de manière absolue le système de la langue.

Notre article consiste à faire la lumière, sur les différentes applications menées par différentes théories linguistiques dans le but de formaliser à l'absolu les systèmes linguistiques mais qui se sont avérées vaines à la fin.

**Nos hypothèses** seront formulées comme suit :

— **La rupture épistémologique qui a marqué l'avènement du structuralisme comme doctrine n'aurait pas été en mesure de structurer de manière formelle et à l'absolu les systèmes linguistiques permettant à des algorithmes (IA) de jouer un rôle prépondérant dans l'enseignement des langues.**

— **L'usage est un facteur prédominant qui empêcherait toute application algorithmique (IA) afin d'arriver à des résultats probants d'un point de vue formel.**

Notre travail consiste donc à interroger certaines théories circonscrites dans le champ de la linguistique et la sociolinguistique américaines<sup>1</sup> que nous compléterons par le point de vue de la sociolinguistique française afin de confirmer ou d'infirmer nos hypothèses.

---

<sup>1</sup>Il s'agit ici du distributionnalisme, du fonctionnalisme et de la théorie générative et transformationnelle que nous aurons à confronter avec la sociolinguistique américaine au travers de la théorie de Dell Hymes/Widdowson et de l'école française au travers de Françoise Gadet à partir d'exemples tirés de son livre : *La variation sociale en France.*

## 1. Du structuralisme américain et de ses applications

### 1.1. La syntaxe distributionnelle

La *théorie distributionnelle*, dont le précurseur fut Léonard Bloomfield<sup>2</sup>, a été la première théorie à poser les jalons d'une première approche formaliste qui a tenté d'élucider le problème des rapports du sens avec la forme linguistique au moyen de l'analyse en constituant immédiat (CI). Cette théorie américaine considérait la langue comme une pure forme et entend la définir en termes qui bannissent toute référence au sens, à la psychologie ou à quelque contenu mental. À ce propos, Bloomfield partant du postulat suivant :

« Tous les mots qui pourraient occuper le même nombre de position dans des énoncés libres minimaux appartiennent à la même classe distributionnelle »

a circonscrit son analyse à un corpus clos – un ensemble déterminé de phrases<sup>3</sup> – et à les réduire à ses éléments qui seront ensuite classés et répertoriés à la façon d'un algorithme<sup>4</sup>, c'est-à-dire, une suite d'opérations rigoureuses, appliquées mécaniquement. Cette analyse a pour objet de réduire le discours à ses constituants immédiats qui seront classés, hiérarchisés au moyen d'opérations formelles telles que : *la substitution*, *le parallélisme*, etc.

L'illustration de cette démarche a été réalisée grâce à la représentation<sup>5</sup> schématique empruntée à Charles Hockett.

Tableau 1 – La boîte de Hockett<sup>6</sup>

elle (3)	abord (10)	-ait (11)	un (6)	problème (8)	délicat (9)
	abordait (4) un problème délicat (5)				
	abordait un problème délicat (2)				
<b>elle abordait un problème délicat (1)</b>					

Dans ce cas de figure, les illustrations tendent à démontrer l'orientation qu'a prise la théorie distributionnelle dans l'étude du système de la langue. Cette étude à caractère structurale qui se voulait formelle et qui procédait par algorithme à l'instar de l'intelligence artificielle n'a pas été toujours rigoureuse dans sa démarche dans la mesure où elle fut confrontée à des limites opérationnelles et aussi à certaines ambiguïtés. Ces limites opérationnelles ont été mises en exergue par le fait qu'en voulant procéder par algorithme, elle ne renseignait pas le locuteur sur les raisons pour lesquelles des phrases ayant des mêmes

<sup>2</sup> Linguiste américain, auteur de l'ouvrage intitulé *Language* dans lequel il développe sa théorie sur la syntaxe distributionnelle. Il s'est investi, dans un premier temps, dans l'étude des langues amérindiennes.

<sup>3</sup> L'objet d'étude de la syntaxe étant dans ce cas, *la phrase*.

<sup>4</sup> Le rapprochement est vite fait entre l'IA suite d'algorithmes et l'approche distributionnelle formelle.

<sup>5</sup> L'exemple cité en référence est tiré de *Clefs pour la linguistique* de G. Mounin (Seghers).

<sup>6</sup> Charles Francis Hockett était un linguiste américain qui a développé de nombreuses idées influentes en linguistique structuraliste américaine. Il représente la phase post-bloomfieldienne du structuralisme, souvent appelée « *distributionalisme* » ou « *structuralisme taxinomique* ».

distributions des mêmes constituants immédiats sont l'une acceptable, à l'exemple de la phrase : « *La postérité admire Ibn Khaldoun* » et l'autre inacceptable : « *La Sphéricité admire Ibn Khaldoun* » ; ou toutes deux correctes, mais ayant deux sens, donc deux syntaxes différentes dans l'optique transformationnelle malgré l'apparence distributionnelle comme dans les exemples (« *Les inondations d'Alger* » # « *Les inondations de la Casbah* »). Par ailleurs, elle ne permet pas non plus de rendre compte facilement de toutes les structures incises ou intercalées (« *Tu l'as troubles, reprit cette bête cruelle, et je sais que de moi tu médis l'an passé* ») uniquement par la considération des distributions.

Cette démonstration apparemment très fondée a poussé la théorie transformationnelle<sup>7</sup> à tenter de résoudre ce problème avec élégance et surtout sans *a priori*.

### 1.2. La syntaxe transformationnelle

Pour le courant transformationnel, deux phrases ayant la même structure distributionnelle, par exemple<sup>8</sup> :

« Le boulanger cuit le pain »  
« Le boulanger cuit la nuit »

n'ont pas la même structure syntaxique lorsqu'elles n'acceptent pas les mêmes transformations. Pour les deux phrases précédentes :

#### Transformation interrogative par que ? (Ou qui ? etc.)

« Que cuit le boulanger ? Le pain. (Possible)  
« Que cuit le boulanger ? La nuit (impossible).

#### Transformation interrogative par quand ?

« Quand cuit le boulanger ? Le pain. (Impossible)  
« Quand cuit le boulanger ? La nuit. (possible).

#### Transformation passive

« Le pain est cuit par le boulanger (possible)  
« La nuit est cuite par le boulanger (impossible).

De ces approches analytiques tendant à tirer la conclusion que toutes les phrases d'apparences identiques, en termes de structure, s'avèrent être différentes lorsqu'on leur applique les opérations de transformation. Par conséquent, la théorie distributionnelle ne pouvait pas répondre à une démarche trop formelle c'est-à-dire algorithmique à l'instar de l'IA.

De ces différentes applications, Harris et Chomsky ont tenté de combler les insuffisances constatées, au niveau des applications de la théorie distributionnelle, pour se

<sup>7</sup> Harris a été le précurseur de cette théorie alors qu'il appartenait auparavant au courant distributionnaliste.

<sup>8</sup> Phrases tirées de *Clefs pour la linguistique* de Georges Mounin.

rapprocher des processus algorithmiques de l'IA en matière de formalisme pure et dure<sup>9</sup>. Ils ont élaboré une théorie selon laquelle la syntaxe est composée d'un stock relativement réduit de phrases de base ou phrases-noyaux, matériau ou nucléaire (*kernel*s) dont toutes les autres phrases de la langue, c'est-à-dire du système, peuvent et doivent être dérivées par des opérations de transformations – à partir du postulat :

« Tout individu est capable de produire et de comprendre des phrases qu'il n'a jamais lu ou entendu auparavant »<sup>10</sup>.

Ce procédé qui tend à se rapprocher d'un processus algorithmique dont l'objectif est de générer toutes les structures d'une langue, d'un système a montré ses limites et, suscité beaucoup de réserves. En effet, si la technique par laquelle les transformations produisent toutes les phrases possibles à partir des *kernel*s (phrases noyaux) veut rester formelle, ce qui est le cas des processus algorithmiques, il lui faut tout expliquer, tout particulièrement, comment le locuteur apprend à distinguer les produits inacceptables des produits acceptables – à l'exemple des phrases suivantes :

1/ *La réponse est arrivée par le courrier et  
2/ le courrier arrive la réponse.*

En termes de structures formelles, nous aurons :

1/ P SN<sub>1</sub>+V + SN<sub>2</sub> > SN<sub>2</sub> +SV [AUX +V+PPE.] + par + SN<sub>1</sub>

Formalisées ainsi, à la manière d'un algorithme, l'une des deux structures n'est pas valide puisqu'elle sera soumise aux contraintes de la sémantique.

Par extrapolation, l'établissement d'une nomenclature de ces hypothétiques compatibilités constituerait un défi presque insurmontable en matière de procédé algorithmique qui pourrait entraver un bon travail d'analyse et de réflexion sur le système linguistique en matière grammaticale.

### 1.3. La théorie générative transformationnelle ou le formalisme à l'absolu

À la suite de cette approche transformationnelle qui s'est trouvée en butte à des obstacles insolubles, N. Chomsky s'est lancé dans la conception d'une théorie formalisée à l'absolu intitulée *grammaire générative et transformationnelle* (GGT). À partir de règles de réécriture traduites par des symboles et des signes, il a démontré, à sa manière, comment les locuteurs d'une langue se construisent progressivement la syntaxe de leur langue au moyen de la manipulation de ces règles de réécriture en pensant aussi également que sa théorie de la syntaxe générative jetterait une lumière nouvelle, logico-métaphysique, sur l'origine du langage. Cette grammaire qui fut indiscutablement la plus ambitieuse et la plus importante des théories récentes sur le langage appelle les mêmes objections que les théories

<sup>9</sup> De manière générale, les résultats obtenus au moyen de l'IA ne doivent pas être remis en cause en matière d'interprétation.

<sup>10</sup> À l'instar de Bloomfield et de son postulat annoncé précédemment, Chomsky a tenté d'asseoir son approche sur un postulat qui traduit l'idée de *créativité propre au sujet parlant*.

précédentes, la syntaxe distributionnelle et la syntaxe transformationnelle puisqu'elle recourt indirectement aux critères sémantiques intuitifs que l'on peut expliquer de la manière suivante :

Les règles de la grammaire générative devront se présenter sous forme d'instructions explicites qui pourraient s'appliquer de façon mécanique, à l'exemple d'un logiciel. À un ensemble d'instructions de cette sorte, décrivant de manière rigoureuse une succession d'opérations élémentaires qui aboutissent à la résolution d'un problème global (en grammaire, la production de phrases), on donne le nom d'*algorithme*<sup>11</sup>.

En ce sens, on peut dire qu'une grammaire générative réussie serait un algorithme. Ce procédé algorithmique s'exprime ainsi :

$$P \rightarrow SN + SV$$

est l'une des premières règles de cette grammaire linguistique.

#### Elle se lit :

la phrase noyau (P) se réécrit par (  $\rightarrow$  ) un syntagme nominal (SN) concaténé à (+) un syntagme verbal (SV).

Dans ce type de grammaire, l'usage des symboles est indispensable d'où l'idée de formalisation à la manière d'un algorithme. Cette théorie formelle, qui tend vers l'explicite en matière de démarche est fondée sur des règles de réécriture dont l'une des particularités est la récursivité<sup>12</sup>. Cette recherche de l'explicite entraîne des descriptions longues que l'on trouve effectivement dans les grammaires scientifiques portant sur des points limités et qui font l'objet de traitement spécifique au moyen de logiciel<sup>13</sup> dans un but purement scientifique. Mais dans une grammaire d'enseignement – pédagogique –, aujourd'hui, il n'est pas possible de réaliser totalement cette ambition.

Afin d'affiner au mieux l'analyse explicite du système linguistique à partir de sa grammaire, Chomsky a circonscrit la tâche la plus élémentaire qu'une grammaire doit remplir :

« une grammaire doit être capable d'énumérer explicitement toutes les phrases qui sont incontestablement grammaticales, ou bien formées, dans la langue étudiée, et d'exclure explicitement toutes les séquences qui sont incontestablement agrammaticales dans cette langue ».

Cela renforce l'idée que la grammaire générative se veut être un mécanisme analogue à une machine à calculer et qui énumère (ou engendre) les phrases grammaticales au moyen d'instructions qui sont l'équivalent de règles grammaticales. Cette idée nous fait penser à un

<sup>11</sup> La GGT et l'IA se partagent jusque-là un même support sous forme de soubassement : le concept d'*algorithme*.

<sup>12</sup> Une règle est dite *récursive* lorsqu'elle s'applique autant de fois que l'on veut.

<sup>13</sup> Noam Chomsky a eu recours à ce procédé analytique à la suite de *l'étude des langues des tribus primitives amazoniennes*.

logiciel qui fonctionne selon des paramètres préétablis. Ce procédé fondé sur des concepts tels que métalangage, métalangue, terminologie, générer, est d'un usage courant en mathématiques.

Voulant renforcer davantage les mécanismes spécifiques à ce type de grammaire, N. Chomsky a explicité le concept de *grammaticalité* dans le but de cerner avec précision les contours de ce concept :

la grammaticalité, dit-il, permet d'éviter les confusions qui résultent de l'emploi de notions vagues comme celles de « phrases possibles » (ou « impossibles »), « correctes » (ou « incorrectes »), « existantes » (ou « inexistantes »).

C'est pourquoi, dans le souci d'asseoir ce concept de la grammaticalité, il le circonscrit aux critères suivants :

### **Phrase grammaticale ne veut pas dire phrase dotée de sens**

Chomsky en donne deux exemples (traduit par Nicolas Ruwet<sup>14</sup>) qui montrent bien la différence entre grammaticalité et signification.

**Séquence 1** : *D'incolores idées vertes dorment furieusement.*

**Séquence 2** : *Furieusement dormir idées vertes incolore.*

De ces deux séquences, toutes deux, également dépourvues de sens, la première possède une structure qui peut la faire admettre pour une phrase. Elle est grammaticale. La seconde n'est qu'un « *tas de mots* », elle est *agrammaticale*.

Par conséquent, si en termes de formalisme, le traitement au moyen de logiciel, accepte la première et rejette la seconde, la contrainte sémantique empêchera de saisir le sens réel de la première bien qu'elle en soit pourvue.

### **Phrase grammaticale ne veut pas dire phrase réelle (relevé dans un corpus).**

Un corpus, si vaste soit-il, ne contient qu'un nombre limité de phrases.

Nul corpus ne peut prétendre fixer la liste des phrases grammaticales. D'autre part, un corpus, dans la mesure où il est l'expression de la performance<sup>15</sup> et non de la compétence, est toujours « plus ou moins accidentel » et comporte des éléments extralinguistiques : variations stylistiques, dialectales, « ratés » à mettre en relation avec la situation d'émission (variations de l'attention, émotion, troubles divers...).

À ce propos, la dichotomie compétence/performance montre clairement la nuance qui puisse exister entre ces deux concepts dans la mesure où la notion de « langue » n'est pas aussi générale comme nous pouvons le supposer.

Par conséquent, si la **compétence** est définie comme « *la connaissance que l'auditeur a de sa langue* », et la **performance** comme « *l'utilisation réelle dans situations concrètes* » de la

<sup>14</sup> La grammaire générative (1966) dans *Langage*, n° 4, Larousse.

<sup>15</sup> L'idée de performance fait basculer le rapport langue/parole/discours vers le concept de l'usage et par conséquent de la diversité linguistique.

compétence<sup>16</sup> ; *la performance ne peut évidemment pas refléter fidèlement la compétence*. Si la compétence, la possession du mécanisme linguistique, est théoriquement la même pour tout le monde, la performance, c'est-à-dire, la mise en marche de ce mécanisme, varie considérablement d'un sujet à l'autre, et est fonction de nombreux facteurs comme l'attention, la fatigue, l'émotivité, mais aussi la situation socio-culturelle du locuteur, le type de conversation qu'il tient, l'endroit, etc. En pratique, cela se traduit par une performance marquée par un plus grand relâchement dans le discours, par des répétitions, des interruptions, par un respect excessif ou un non-respect des règles normatives, par le choix non seulement d'un vocabulaire adapté à la situation, mais aussi d'une syntaxe particulière, etc. Cette diversité linguistique en matière de pratique de la langue, entraverait, à coup sûr, l'élaboration d'une grammaire particulière dont la conception pourrait être envisagée de manière uniforme et homogène et présentée en tant que telle à un logiciel en vue d'un traitement particulier.

## 2. De la conception taxinomique à la conception théorique

En effet, il est connu que l'évolution générale de la science est toujours marquée par deux étapes, deux conceptions radicalement différentes de la recherche. La première conception est une conception « taxinomique ». Elle consiste à observer et à classer<sup>17</sup> des faits. Celle-ci nous a permis d'enregistrer une masse énorme de renseignements sur le monde qui nous entoure. La seconde, beaucoup plus ambitieuse, veut non seulement décrire les faits mais les expliquer. Ce que l'intelligence artificielle est loin de pouvoir concrétiser au niveau de la langue.

Pour la science moderne, disait Nicolas Ruwet :

« il s'agit beaucoup moins de collectionner et de classer des faits que de construire, – à partir d'un nombre limité d'observations ou d'expériences – des théories générales, des modèles hypothétiques, destinés à expliquer les faits connus et à en prévoir de nouveaux » (1967, p. 12).

Cette deuxième conception ne peut être l'apanage des études sur la langue mais plutôt elle n'est le fait que des sciences les plus abouties et les plus avancées. C'est en particulier le cas des sciences dures que nous avons évoquées tout au début de cet article (*les sciences physiques, les mathématiques, la chimie*). Les scientifiques ne se contentent pas d'observer le monde, mais ils émettent des « *modèles hypothétiques* » afin d'essayer d'expliquer les lois qui régissent le monde.

Toujours à propos de cette taxonomie qui consiste à observer et à classer, il n'est pas suffisant pour une grammaire générative et transformationnelle de rendre compte des catégories d'éléments (*noms, verbes, adjectifs*, etc.) ; encore devra-t-elle distinguer des classes telles que, pour les noms : *commun/propre, masculin/féminin, animé/non animé, comptable/non comptable, concret/abstrait*, etc.

<sup>16</sup> Voir Chomsky ([1965] 1971).

<sup>17</sup> Il nous suffit, à ce propos, de relater le phénomène de l'évolution de la classification (ou *catégorisation des mots*) depuis l'Antiquité jusqu'à nos jours.

Par conséquent, les réflexions qui, jusqu'à ces dernières années, ont été faites sur la langue relevaient de la première conception. Les grammaires traditionnelles et structurales étaient toutes des modèles taxonomiques de la langue, c'est-à-dire des modèles préétablis et, de ce fait, pouvaient toutes se ramener à des collections d'exemples plus ou moins bien décrits, à des listes de phrases françaises définies selon des critères « aléatoires » que certaines grammaires structurales transformationnelles, à l'image de celle de Maurice Gross, utilisent sous la forme de symboles et abréviations pour relativiser les approches en matière d'analyse comme par exemple :

- ? Forme difficilement acceptable pour certains ;
- \*? Forme très difficilement acceptable ;
- \* Forme inacceptable.

Cette approche, en somme, tout au moins relative, pourrait engendrer, en termes systémiques, des confusions dans les interprétations, à l'exemple des phrases suivantes où le verbe *appeler* susceptible d'être suivi de deux compléments directs est le pivot de la phrase :

- 1/ « *Soudain quelqu'un appelle l'éclairé Kacem* »<sup>18</sup> étant considérée comme la forme la plus répandue, donc acceptable pour la majorité des locuteurs ; une minorité récalcitrante pourra considérer comme acceptable la phrase :
- 2/ \*? « *Soudain appelle quelqu'un l'éclairé* ».

Cependant si l'on applique, à la structure initiale, la permutation du sujet, les phrases produites ne seront acceptables pour personne, et, par conséquent, étoilées sans point d'interrogation.

- 3/ \* *Soudain appelle quelqu'un l'éclairé « Kacem »* ;
- 4/ \* *Soudain appelle l'éclairé quelqu'un « Kacem »* ;
- 5/ \* *Soudain appelle l'éclairé « Kacem » quelqu'un*.

À l'exception des phrases surmontées d'un astérisque et considérées comme totalement agrammaticales, le critère d'acceptabilité sera tout à fait relatif et nuancé dans le cas de la phrase (2). Les structurelles formelles telles que définies au niveau de la langue seraient sujettes à différentes constructions et conceptions au niveau du discours et, par conséquent, interprétées selon des paramètres préétablis (*situation de communication, intentionnalité des locuteurs, implicites, etc.*)

De fait, ces critères d'acceptabilité rendent le traitement logiciel des structures syntaxiques aléatoires puisque les réponses ne seront pas tout à fait fondées.

Cette nouvelle conception de la langue en termes de structures formelles (*règles de réécritures*) que la théorie chomskyenne a tenté d'imposer fut remise en cause par les nouvelles

<sup>18</sup> Pour illustration, ces exemples ont été tirés du livre *Exercices de syntaxe transformationnelle* de A. Borillo, J. Tamine et F. Soublin.

orientations de certaines recherches à l'image des travaux menés par des sociolinguistes comme Dell Hymes<sup>19</sup> ou Françoise Gadet<sup>20</sup>.

Hymes, à la suite des travaux de Chomsky, n'a pas manqué de réagir contre ce qu'il considérait comme une conception idéaliste de la théorie chomskyenne pour qui la compétence linguistique est la capacité innée que possède un locuteur-auditeur idéal de comprendre et de produire des énoncés jamais entendus auparavant. Hymes reproche surtout à Chomsky de ne pas tenir compte des situations de communication concrètes dans lesquelles toute langue s'utilise, c'est-à-dire des règles d'emploi de la langue qui varient précisément selon les situations et qu'aucun traitement logiciel ne pourrait circonscrire ou délimiter. Afin de bien mettre en évidence cette perspective sociolinguistique, Hymes propose de recourir au concept de compétence de communication qui va au-delà de la compétence linguistique.

Cette nouvelle orientation en matière de didactique des langues a créé un bouleversement sans précédent au niveau de la recherche fondamentale à l'époque. Ce bouleversement a entraîné un foisonnement d'idées qui allaient également influencer le domaine de la didactique des langues. Parmi ces idées, nous citons les paramètres qui circonscrivent la notion de *compétence de communication* et avec lesquels que la didactique devrait compter :

- **une compétence grammaticale** que l'on peut désigner comme la connaissance du code linguistique ;
- **une connaissance sociolinguistique**, comprenant la connaissance des règles socioculturelles<sup>21</sup>
- **une composante discursive** comprenant les règles de cohésion et de cohérence ;
- **une composante stratégique** comprenant les mécanismes de compréhension qu'utilisent les locuteurs pour contourner les ratés de la communication.

Ce curriculum multidimensionnel posera d'emblée la problématique de la diversité linguistique à envisager dans le traitement du savoir, de la norme de la langue à envisager aussi à l'oral et à l'écrit, de l'environnement linguistique, de l'ensemble des éléments non linguistiques qui conditionnent toute communication et qui pourrait peser sur la structure même du message et sur sa compréhension :

- le moment et le lieu de parole,
- les présupposés psychologiques, culturels, historiques, etc. ;

tout un contexte situationnel difficile d'être traité par un logiciel dans le but de rendre le discours intelligible et auquel souvent celui-ci fait référence.

---

<sup>19</sup> Dell Hathaway Hymes est un sociolinguiste, anthropologue et folkloriste américain dont le travail porte essentiellement sur les langues amérindiennes du Nord-Ouest Pacifique.

<sup>20</sup> Françoise Gadet est une sociolinguiste française, née en 1947. Elle est professeure émérite de sociolinguistique à l'université Paris-Nanterre et chercheuse au sein du laboratoire *Modèles, Dynamiques, Corpus*.

<sup>21</sup> À ce propos, nous disons que la formation en didactique des langues, à l'université, est totalement démunie de spécialités dont l'impact est considérable pour promouvoir une formation de qualité dans le domaine socioculturel, nous pensons à *l'histoire des mentalités*, de *l'ethnographie de la communication* ou encore de *l'anthropologie culturelle*.

Dans le même ordre d'idées, nous ne manquerons pas de citer l'exemple de Françoise Gadet en référence au professeur de langue de FLE qui, dans son enseignement axé sur la communication, sait qu'il doit enseigner un registre qui, certes est toujours vivant et utilisé, mais qui est concurrencé sinon menacé par des variantes tout aussi vivantes que le français des natifs.

Cet exemple<sup>22</sup> porte sur l'emploi des nombreuses variantes de l'interrogation partielle « quand vous venez ? » telles qu'elles ont été recueillies :

1. Quand venez-vous ?
2. Quand est-ce que vous venez ?
3. Vous venez quand ?
4. Quand vous venez ?
5. Quand que vous venez ?
6. Quand c'est que vous venez ?
7. Quand que c'est que vous venez ?
8. Quand c'est que c'est que vous venez ?
9. Quand que c'est que c'est que vous venez ?
10. C'est quand que vous venez ?
11. C'est quand c'est que vous venez ?
12. C'est quand que c'est que vous venez ?
13. Vous venez quand est-ce ?
14. Vous venez quand ça ?

Comme cela apparaît dans ce corpus, la diversité des cas pour une même variante rend la tâche plus complexe en matière de *sélection*, de *traitement logiciel* et d'*usage*. Finalement, la chose est fort délicate :

- Quelle(s) structure(s) choisir ?
- Faut-il les accepter toutes ou opérer une sélection ?
- Si oui, selon quels critères ?

Si dans les pratiques langagières, les usages peuvent varier, en milieu scolaire, la norme scolaire impose des règles qu'il faut respecter. Cependant, il existe des situations qui transgressent la norme. Pour répondre à des situations précises, certains linguistes ont cru possible de récupérer le mot « *norme* », et de l'utiliser dans un sens nouveau, où il ne sert plus à distinguer un usage particulier de la langue mais plutôt des usages : *norme scolaire*/*norme sociale*, *norme écrite*/*norme orale*.

Dès lors, un traitement logiciel ne serait pas en mesure de rendre compte de toutes ces situations dans lesquelles les variantes sont légitimes puisque chaque cas est tributaire de critères particuliers comme les types de situation de communication ou encore les conditions sociales des usagers, etc. Le recours, dans ce cas, à la notion de degré d'acceptabilité rend les jugements aléatoires dans la mesure où ces jugements deviennent relatifs.

<sup>22</sup> Cet exemple est emprunté à Françoise Gadet, *La variation sociale en français* (2007).

## Conclusion

En choisissant comme domaine d'application les syntaxes structurales et les travaux qui ont été effectués au gré des options et des difficultés rencontrées en matière de traitement du savoir, nous avons voulu montrer que les différentes solutions proposées dans l'une ou l'autre de ces grammaires demeurent insuffisantes même si le traitement réservé au domaine des connaissances à acquérir en matière d'enseignement apprentissage des langues s'appuyait et s'appuie encore sur les nouvelles technologies telles que l'intelligence artificielle. Les recherches menées dans ce domaine ont toutes été soumises à un traitement assisté par ordinateur<sup>23</sup> mais n'ont abouti qu'à rendre la tâche plus ardue et souvent inextricable.

De prime abord, il nous faut distinguer entre le domaine de la linguistique dont le sous-bassement est *la recherche scientifique* et celui de l'enseignement-apprentissage des langues dont l'objectif est de *faire acquérir des connaissances en cherchant à développer des compétences*.

L'abbé Rousselot<sup>24</sup>, célèbre phonéticien qui, en son temps, est arrivé à identifier sept variantes de la lettre « a » dans le domaine de la phonétique expérimentale, ne s'est jamais mis dans l'idée de les mettre au service de l'enseignement du système phonologique français mais plutôt s'est contenté de les mettre en exergue dans le cadre de son travail de chercheur. Encore que, faut-il, au commun des mortels, saisir les nuances entre ces sept « a ».

Ce qui n'est guère le cas d'un didacticien encore plus d'un enseignant. Pour ce dernier, un des problèmes majeurs qui se pose, dès lors qu'il est question de grammaire, est celui de la place que doit occuper celle-ci dans un programme ou dans un cours de LE par exemple – *Quelle importance faut-il lui accorder dans un cours de LE, en milieu scolaire ou universitaire ?*

Nous pouvons encore ajouter à ce type de questionnement d'autres encore dont la didactique des langues ne peut s'en passer et qui porte sur *les styles d'apprentissage, les besoins langagiers, l'âge des apprenants, leur niveau de connaissance, leur niveau de scolarité, l'hétérogénéité<sup>25</sup> de la classe, les objectifs à atteindre, etc.*

Enfin, si *l'Intelligence Artificielle* est en mesure d'apporter des réponses relatives<sup>26</sup> dans le traitement du savoir en matière de langue, *elle ne peut, en aucun cas, définir une démarche uniforme qui homogénéiserait entre le concept de grammaire d'un côté, la norme ou les normes qui fixent l'usage de la langue à acquérir et les règles qui le définissent*. Les exemples que nous avons cités dans ce modeste travail de réflexion montrent combien les situations sont nombreuses et parfois même difficilement acceptables et interprétables.

<sup>23</sup> Nous citons le cas des traitements assistés par ordinateur menés par N. Chomsky et toute l'équipe du MIT.

<sup>24</sup> Fondateur de la phonétique expérimentale.

<sup>25</sup> Concept mis à jour et soumis tout récemment à débat par J.-L. Chiss et J. David dans *Didactique du français et étude de la langue* (2012, Armand Colin).

<sup>26</sup> Nous venons de voir, dans ce modeste travail de réflexion, qu'aucune théorie n'a donné de réponse absolue aux questions posées.

## Références

- BLOOMFIELD, Leonard (1926). A Set of Postulates for the Science of language. *Language*, N° 2, p. 153-164.
- BORILLO, Andrée ; Joëlle TAMINE ; Françoise SOUBLIN (1981). *Exercices de syntaxe transformationnelle du français*. Armand Colin, coll. Linguistique.
- CALVÉ, Pierre. (1994). « Comment faire de la grammaire sans trahir le discours : cas des exercices grammaticaux ». *Revue canadienne des langues vivantes*, vol. 50 n° 4, p. 636-645.
- CHOMSKY, Noam (1965). *Aspects de la théorie syntaxique* (trad. Seuil, 1971).  
— (1957). *Structures syntaxiques* (trad. Seuil, 1969).
- GADET, Françoise (2007). *La Variation sociale en français*. Paris : Éditions OPHRYS , coll. « L'essentiel français ».
- GERMAIN, Claude ; Hubert SEGUIN (1998). *Le point sur la grammaire*. CLE International.
- GUIRAUD, Pierre (1974). *La grammaire*. Que sais-je ? – Presses Universitaires de France.
- HARRIS, Zellig S. (1954). Distributional Structure. *Word*, Vol. 10, N° 2-3, p. 146-162 (trad. Française, *Langages*, n° 20, 1970).
- HYMES, Dell H. (1972). « On communicative compétence ». Dans J. B. PRIDE ; J. Holmes (Eds.) (1972). *Sociolinguistics: Selected Readings* (p. 269-293). Penguin Books.
- LAGARDE, Jean-Pierre (1988). « Les parties du discours dans la linguistique moderne et contemporaine ». *Langages*, n° 92, p. 93-108.
- LE SOIR d'Algérie (2025, 24 avril). Supplément IA : « Le soir intelligence artificielle ».  
— (2025, 17 avril). Supplément IA : « Le soir intelligence artificielle ».  
— (2025, 16 jan.). Supplément IA : « Le soir intelligence artificielle ».
- MITCHELL, Rosamond (1994). Foreign Language Teachers and the Teaching of Grammar. Dans Martin BYGATE ; Alan TONKYN, & Eddie WILLIAMS (Eds.) (1994). *Grammar and the Language Teacher* (p. 215-223). Hemel Hempstead: Prentice Hall.
- MOUNIN, Georges (1978). *Clefs pour la linguistique*. Seghers.
- NIQUE, Christian (1991). *Initiation méthodique à la grammaire générative*. Armand Colin, coll. Linguistique, 2<sup>e</sup> éd.
- NUNAN, David (1994). « Linguistic Theory and Pedagogic Practice ». Dans Terence ODLIN, (éd.) (1994). *Perspectives on Pedagogical Grammar* (p. 253-270). Cambridge Applied Linguistics. Cambridge University Press.
- RUWET, Nicolas (1972). *Théorie syntaxique et syntaxe du français*. Paris : Éditions du Seuil.
- WIDDOWSON, Henry G. (1980). « Applied Linguistics: The Pursuit of Relevance ». Dans Robert B. KAPLAN (éd.). *On the Scope of Applied Linguistics* (p. 74-84). Newbury House.

### Pour citer cet article

Aicha Ibtissem  
BOUACHA,  
« L'Intelligence Artificielle  
et l'étude du système  
linguistique :  
Entre contraintes et  
facilités d'acquisition des  
connaissances »,  
*Paradigmes*, vol. IX, n° 02,  
mai 2026, p. 253-265.