PRÉSENTATION DU COURS

Ce semestre sera pour l'essentiel consacré à l'étude de la forme des mots, c'est-à-dire à la morphologie, mais auparavant, nous allons consacrer quelques séances à un retour sur l'étude de la matière phonique du français. Nous allons reprendre, sur la base de ce que vous avez vu en 1re année en linguistique-phonétique, les rapports entre l'oral et l'écrit, et puis nous nous intéresserons à la phonologie, qui a pour objet l'étude de la fonction des sons à l'intérieur du système de la langue. Nous pourrons ainsi comprendre comment les deux niveaux d'analyse, phonétique et morphologique, se complètent. Nous verrons par exemple que la formation de certains morphèmes est tributaire de l'environnement phonologique dans lequel il se trouve : par exemple la variation du morphème de pluriel en français, qui s'adapte à la configuration phonologique finale du radical : il se réalise par **[z]** (liaison) après une voyelle (*les amis* /le.z‿a.mi/) mais par l'allomorphe zéro (noté ∅) après une consonne, qu'elle soit sonore ou sourde (*les chats* /le∅ ʃa/, *les dons* → /le∅ dɔ̃/). Ou encore le morphème préfixal ***dé-***, qui possède deux allomorphes (***dé-*** et ***dés-***) dont la distribution est régie par la nature phonologique de l'initiale de la base lexicale (consonne ou voyelle : *faire* – ***dé****faire*, *accorder* – ***dés****accorder*).

CODE ORAL, CODE ÉCRIT

Pour travailler sur l'oral des différentes langues, on dispose de l'Alphabet Phonétique International (A.P.I.). La première version de l'API a été publiée en 1888 par l'Association phonétique internationale, sous la direction du linguiste français Paul Passy, puis il a été révisé plusieurs fois au cours du XXe siècle. L'objectif de l'API consiste à transcrire l'oral de toutes les langues en utilisant des symboles universels. L'oral est transcrit non pas en passant par les contraintes de l'orthographe, mais tel qu'il se prononce. Le principe de l'API part d'un principe simple : à chaque son on fait correspondre un signe et un seul, et à chaque signe, un son et un seul. C'est ainsi qu'en français on compte 36 sons : 16 voyelles (dont 3 ont un statut problématique : [ə], [œ], [ɑ]), 17 consonnes et 3 semi-voyelles, ou glides ([j], [ɥ] et [w]).

Pour l'écrit, on dispose d'un alphabet de 26 lettres, censées représenter graphiquement les 36 sons du français. C'est un alphabet issu de l'alphabet latin, lui-même issu de l'alphabet grec. Le principe de l'écriture alphabétique est d'associer des signes graphiques à des sons, contrairement à d'autres systèmes de transcription, comme l'écriture idéographique, qui transcrit des idées, des concepts. Mais l'alphabet latin est relativement inadapté aux sons du français, ce qui a conduit à des combinaisons de lettres pour transcrire des sons, par exemple V+*n* pour les voyelles nasales (*an*, *en*, *on*, *in*, etc.), et à l'ajout de lettres permettant de transcrire des mots étrangers, comme *w*, (*wagon*, *week-end*, etc.)

Commenté [ST1]: Le (e) muet

Commenté [ST2]: Voyelle nasale ouverte antérieure arrondie, que l'on distingue mal de **[ɛ̃]**, comme dans *brun* **[brɛ̃]** et *parfum* **[parfœ̃].**

Commenté [ST3]: Voyelle postérieure et ouverte, différente de **[a]** antérieur mais dont la réalisation en français standard n'est plus ressentie.

Commenté [ST4]: C'est le même son que l'on retrouve dans le mot anglais « Yes ». Le [j] se réalise à l'écrit par plusieurs configurations graphiques : il peut s'écrire (y) comme dans (yeux), (i) comme dans (piano), (ï) comme dans (aïeux), (il) en fin de mots comme dans (travail), (ill) comme dans (feuille).

Commenté [ST5]: Le son [ɥ] s'écrit d'une seule manière : la lettre (u) suivante d'une autre voyelle comme dans (huit), (buée), etc.

Commenté [ST6]: Le son [w] se réalise par sept graphies : (oi) comme dans (boire), (ou) comme dans (oui), (oin) comme dans (soin), (qu) comme dans (adéquat), (gu) comme dans (jaguar), (w) comme dans (watt), (oë) comme dans (poêle), et (oe) comme dans (moelle).

ou de nouveaux sons du français, comme *j* et *v* (lettres ramistes, XVI[e] siècle). De plus, avec le temps, le système d'écriture s'est éloigné de la réalité sonore de la langue, car l'écrit est plus fixe que l'oral. Du coup, il n'y a plus (ou pas) de correspondance numérique entre sons et lettres, ce qui implique qu'une lettre va être sollicitée pour transcrire plusieurs sons. Cette dysharmonie est encore aggravée si l'on compare la transcription orthographique des voyelles et des consonnes. On a effet 6 signes graphiques vocaliques (*aeiouy*) pour transcrire 16 sons vocaliques, et 20 signes graphiques consonantiques pour transcrire 17 sons consonantiques. On ne dispose pas, de plus, de lettre spécifique pour noter les semi-voyelles.

Commenté [ST7]: Les lettres ramistes sont les lettres **J** et **V** de l'alphabet latin, distinctes respectivement de **I** et **U**. Ce terme fait référence à l'humaniste **Pierre de La Ramée** (ou Petrus Ramus), qui a proposé en **1562** dans sa *Gramere* d'introduire ces deux lettres pour noter les valeurs consonantiques des semi-voyelles [j] et [w], alors que I et U notaient les voyelles [i] et [u].

Sons	**Lettres**
16 sons vocaliques	6 signes vocaliques
17 sons consonantiques	20 signes consonantiques
3 glides	Rien

On a donc des distorsions entre code oral et code écrit, qui tiennent à l'histoire de la langue. Périodiquement, au gré des préoccupations ministérielles, émerge la question de la réforme de l'orthographe. Trois positions sont en conflit :
— les traditionalistes veulent préserver l'héritage ancestral : pour eux, les aberrations orthographiques (cf. le pluriel de *ail* : *aulx*) sont les plus sûrs atouts du français.
— les réformateurs souhaitent rendre plus *fonctionnelle* l'orthographe française. Par exemple, le *s* de *temps* est la marque du cas-sujet dans l'ancien français, qui a perdu depuis fort longtemps toute fonctionnalité. Ne pourrait-on pas rendre facultative sa graphie ? Ce qu'ils ont prôné dernièrement (1991) : l'abolition du trait d'union (soudure des mots français : auto-stoppeur, autostoppeur) ou celle de l'accent circonflexe, excepté pour les cas ambigus : sur / sûr, le passé simple.
— les révolutionnaires n'ont pas droit de cité ou uniquement dans une visée ironique ou ludique (comme dans la première phrase du roman *Zazie dans le métro* de Raymond Queneau : « doukipudonktan (d'où qu'il pue donc tant ? », où l'auteur supprime les espaces blancs entre les mots »).

Malgré les apparences, l'orthographe du français n'est pas anarchique, mais organisée en un système, qui a été décrit par la linguiste Nina Catach (elle parle plutôt de plurisystème, ou système de systèmes). On se sert, pour décrire ce système, non pas de l'unité graphique concrète, de la lettre, puisque, comme on l'a vu, la correspondance entre lettres et sons est problématique, mais d'une unité graphique abstraite qu'on appelle le graphème. Le graphème est la plus petite unité distinctive de la chaîne écrite ; cette unité peut être composée d'une lettre unique ou d'un groupe de lettres (digramme : 2 lettres ; trigramme : 3 lettres ; tétragramme : 4 lettres), d'une lettre accentuée ou pourvue d'un signe

auxiliaire (la cédille). Les graphèmes, qui sont autant de façons d'écrire les sons, sont beaucoup plus nombreux que ces derniers : pour écrire le son [o], on pourra ainsi utiliser la lettre *o*, la lettre accentuée *ô*, mais aussi les digrammes *ot*, *au*, ou le trigramme *eau*. Par exemple le mot *grammaire* est composé de 6 sons [gramɛr], de 9 lettres : g,r,a,m,m,a,i,r,e et de 7 graphèmes : g,r,a,mm,ai,r,e.
Le graphème a en premier lieu une valeur phonique, puisqu'au départ, l'alphabet retranscrit des sons ; mais il peut aussi avoir une valeur sémantique ou grammaticale. On compte trois principales fonctions des graphèmes, et ainsi trois catégories de graphèmes. On distingue les **phonogrammes**, les **morphogrammes** et les **logogrammes**.

- Les **phonogrammes** sont les graphèmes qui ont une fonction phonétique, c'est-à-dire qu'ils sont chargés de transcrire les sons, d'assurer une représentation graphique directe des unités sonores. Ces graphèmes phonogrammiques sont, dans la plupart des cas, prononcés. Ils représentent donc la partie prononcée de l'orthographe, et constituent 80 à 85 % des unités graphiques. Mais la correspondance n'est pas totale entre lettres et sons, c'est-à-dire qu'il n'y a pas équivalence un son = une lettre. Par exemple, dans le verbe *acquitter* [akite], si le son [a] est bien représenté graphiquement par une lettre unique *a*, c'est un trigramme (*cqu*) qui réalise le son [k], et des digrammes qui réalisent les sons [t] (digramme *tt*) et [e] (digramme *er*). On a donc des correspondances entre des graphèmes (lettres uniques, di- ou trigrammes) et des sons. Les correspondances sont relativement fixes, par exemple le digramme *ou* représente le son [u], le digramme *on* le son [ɔ̃], la lettre à signe auxiliaire *ç* le son [s], le trigramme *ein* le son [ɛ̃], etc. Il faut cependant noter que cette correspondance peut varier en fonction de la position des graphèmes les uns par rapport aux autres. Par exemple, la lettre *s* correspond généralement au son [s], comme dans *valse*, mais au son [z] lorsqu'elle est entre deux voyelles, comme dans *poison*. Si l'on veut rétablir le son [s], il faut alors recourir au digramme *ss*, comme dans *poisson*. De même, la lettre *m* correspond généralement au son [m], comme dans *ami*, mais lorsqu'elle suit une voyelle et qu'elle est suivie par un *b* ou un *p*, elle constitue alors avec la voyelle un digramme correspondant à une nasale **[ɑ̃], [ɛ̃], [ɔ̃],** comme dans *ampli*, *embaumer*, *limbes* ou *composer*.

- Les **morphogrammes** sont les graphèmes qui ont une fonction morphologique. Ils ont pour fonction de transcrire les morphèmes, en particulier les morphèmes grammaticaux. Ce sont eux qui assurent les indications grammaticales, les désinences qui s'ajoutent aux mots pour apporter des renseignements sur le nombre, le genre, la personne, etc. Dans ce cas,

(graphèmes donnant des informations flexionnelles), on parle de morphogrammes grammaticaux. Ce sont, par exemple, les marques de genre et de nombre sur des noms ou des adjectifs, qui sont apportées par les graphèmes *e* et *s* : *petit*, *petit**e*** *; petit*, *petit**s***. Ou encore, les marques de mode, de temps, de personne, de nombre, apportées à des verbes par certains graphèmes, comme dans *je voi**s** / que je voi**e*** (marques respectives de l'indicatif et du subjonctif) ; *je marche / tu marche**s** / ils marche**nt*** (marques de la 2e personne, du pluriel) ; *nous marchons / nous march**i**ons* (marque de l'imparfait). Évidemment, une même lettre (comme le *s*) peut assurer plusieurs fonctions morphogrammiques différentes : pluriel, 2e personne. Ces morphogrammes ne sont pas toujours prononcés (contrairement aux phonogrammes). Par exemple, le morphogramme *t* de 3e personne verbale est non prononcé dans *il vient* : [ilvjɛ̃] et prononcé dans *vient-il ?* : [vjɛ̃til]. Les morphogrammes représentent de 3 à 6 % des unités graphiques, et les plus importants sont des morphogrammes grammaticaux, comme les lettres *s* (ou *x*), *t* (ou *d*), *e*, *r*, ou le digramme *nt*. Il existe aussi des morphogrammes dit lexicaux, qui marquent la relation d'un mot à ses dérivés, à sa famille de mots. Ainsi, le *d* de *lard* le relie-t-il à *larder* et à *lardon*. Mais leur organisation est moins régulière que celle des morphogrammes grammaticaux.

- Les **logogrammes**, enfin, sont des graphèmes, dépourvus de réalisation sonore (ils ne se prononcent pas), qui ont pour fonction exclusive de distinguer dans l'écriture des formes homophones. Les logogrammes jouent ainsi un rôle sémantique, puisqu'ils permettent de distinguer des homophones. Par exemple, dans les couples *a / à*, *croit / croît*, *des / dés*, *mur / mûr*, etc., l'accent a une fonction logogrammique. Le doublement des consonnes (*date* / *datte*), l'utilisation de lettres muettes (*h* et *e*) ont la même fonction logogrammique (*hauteur* /*auteur*, *hôte* / *ôte*, *heur* / *heure*, *ni* / *nie*, *tu* / *tue*).

Ce sont là les trois niveaux, les trois fonctions des graphèmes. Il faut noter cependant qu'il y a parfois des lettres qui ne répondent à aucune de ces fonctions, mais qui sont des sortes de reliquats étymologiques, ou historiques. Par exemple, le *h* de *rhétorique* ou de *théorie* n'a de fonction ni phonogrammique (il ne se prononce pas), ni morphogrammique (il ne marque pas de genre ou de nombre), ni logogrammique (il ne sert pas à distinguer les mots d'éventuels **rétorique* et **téorie*). Leur présence se justifie par la transcription des lettres grecques thêta (th) et rho avec esprit rude (rh).
On peut revenir à présent sur la notion de plurisystème, de système de système. En effet, l'orthographe française superpose trois systèmes qui se recoupent et s'imbriquent : les phonogrammes sont, plus ou moins exactement, en relation avec l'oral, les sons, et ce système semble prioritaire. Les morphogrammes sont en relation avec la grammaire, pour dire vite, dans un système encore assez régulier. Enfin, les logogrammes ne concernent que des couples, des séries lexicales limitées. L'orthographe française est le résultat de ces trois systèmes :

les **phonogrammes** et les **logogrammes** correspondent à ce qui était traditionnellement nommée orthographe d'*usage* ou de mots ; les **morphogrammes** à l'orthographe d'*accord*.

Il faut aussi bien comprendre qu'il n'y a pas des graphèmes strictement phonogrammiques, ou logogrammiques, mais qu'un même graphème peut, suivant son emploi, jouer plusieurs rôles. Par exemple le graphème *s* peut être phonogrammique, représentant le son [s] en alternance avec *ss* à l'intervocalique : *rose*/*rosse*, et en concurrence avec les graphèmes *c*, *ç* : *sa/ça*, mais aussi morphogrammique, marquant le pluriel des noms et des adjectifs, ou la 2eme personne : *maisons*, *fous*, *tu as* vs *il a*, et enfin logogrammique dans le couple : *moi/mois*. De même, si on prend le couple *chant / champ*, la lettre finale, qui n'a pas de fonction phonogrammique, puisqu'elle n'est pas prononcée, a une fonction morphogrammique, car elle marque le lien de ces mots avec certains de leurs dérivés (*chan**t**eur*, *cham**p**être*), et une fonction logogrammique de distinction entre ces deux homophones.

Exercice :

Étudier la correspondance oral vs écrit dans ce passage, puis dégager les fonctions des graphèmes :

Madame Archer est impressionnante, je devine ses muscles sous le coton élastique de son caleçon, elle est possédée par son récit, je voudrais savoir le prénom de sa sœur, elle répète ma sœur, et mon beau-frère et le vin, et les coups, elle décrit minutieusement les traces de coups sur une personne sans nom. Elle a posé les draps que nous nous apprêtions à étendre, elle les a posés sur un tabouret rond recouvert d'une dentelle grise.

(G. Brisac, *Pour qui vous prenez-vous ?*)

Corrigé

On remarque d'abord que l'équivalence stricte entre sons et lettres constitue une exception : rares sont les mots qui présentent cette correspondance (*sur*, *le*, *je*, *a*, *posé*). Si on ne tient pas compte de la graphie du *e* muet, cette correspondance est un peu plus fréquente (*devine*, *grise*, *d'une*).

Fonction phonogrammique : c'est le rôle de l'essentiel des graphèmes. La correspondance graphème-son est bien mieux établie : *c-o-t-on* = [kotɔ̃], *é-l-a-s-t-i-que* = [elastik]. Bien sûr, les graphèmes peuvent être constitués de lettres seules, de digrammes (les classiques *on* ou *om* pour [ɔ̃], *ai* pour [ɛ]), de trigrammes (*ous* pour [u]). On remarque aussi que plusieurs graphèmes peuvent être utilisés pour transcrire un même son (le son [e] est transcrit par *est*, *(s)es*, *é(lastique)*, *et* ; le son [s] est rendu par *s* dans *muscles* ou *élastique*, par *ss* dans

impressionnante, *possédée*, par *ç* dans *caleçon*, par *c* dans *récit*, par *t* dans *minutieusement*). Inversement, le même graphème peut rendre des sons différents, comme *t*, qui peut rendre le son [s], comme on vient de le voir, et comme c'est fréquent dans les terminaisons en *–tion* (*attention*, *mention*…), mais aussi, bien sûr le son [t], comme dans *tabouret*.

Fonction morphogrammique : on reconnaît des morphogrammes grammaticaux dans toutes les marques de flexion, nominale (*s* du pluriel dans *muscles*, *coups*, *draps*, *e* du féminin dans *impressionnante*, *grise*) ou verbale (*e* du féminin dans *possédée*, *s* du pluriel dans *posés*, *r* du conditionnel et *ais* de la 1re personne dans *voudrais*, *i* de l'imparfait et *ons* de la 4e personne dans *apprêtions*). Il y a aussi des morphogrammes lexicaux, comme le *t* final de *récit*, qui le rattache à *réciter*, *récitation*, le *d* final de *rond* et de *drap*, qui les rattache à *rondeur*, *rondement*, *arrondir*, et à *draper*, *drapier*, *draperie*, le *m* final de *prénom* qui le rattache à *nommer*, *prénommer*, etc.
En revanche, on ne peut pas voir de morphogramme dans le *p* final de *coup*, qui ne le rattache pas, dans notre perception moderne de la langue, à *couper*, *coupure*, etc., même si l'origine étymologique est la même. On sera plutôt tenté de voir dans ce *coup* un logogramme, qui permet de distinguer ce mot de ses homophones *cou* et *coût*.