

## Chapitre 3

Syllabe et légalité phonotactique

## SYLLABE ET LEGALITE PHONOTACTIQUE

Chaque langue du monde présente des caractéristiques qui la distinguent des autres langues. Dans le cas de la structure sonore de la langue, objet d'étude de la phonologie, ces spécificités concernent aussi bien l'inventaire des sons individuels que celui des séquences de sons qui sont attestées dans cette langue. Les théories développées en phonologie ont pour objet de décrire les mécanismes qui permettent de transformer une représentation phonologique dite sous-jacente (représentation abstraite de la séquence de phonèmes à prononcer) en une forme de surface (la forme phonétique) correspondant à ce qui est effectivement prononcé. Les modèles proposés définissent donc deux composantes essentielles : la forme des représentations sous-jacentes et les procédures permettant de transformer ces représentations en forme phonétique de surface. Ces règles peuvent être universelles -elles existent dans toutes les langues- ou paramétriques -auquel cas seules certaines langues en sont pourvues. Les représentations phonologiques et les règles qui sont décrites en phonologie peuvent être assimilées à un ensemble de 'connaissances implicites' qui serait stocké par le système cognitif du locuteur (Pierrehumbert, 1990). Cette composante phonologique serait donc partie intégrante d'une base de données -une grammaire- qui serait construite par l'être humain au cours de son développement. Ceci entrouvre la possibilité de proposer une relation étroite entre la grammaire

décrite en phonologie et les processus de perception de la parole étudiés par les psychologues cognitivistes.

Nous avons vu dans le chapitre précédent que des données expérimentales obtenues avec les tâches de détection de phonèmes (Vroomen & de Gelder, 1999) et de *word-spotting* (McQueen, 1998) ont été interprétées comme reflétant l'utilisation par le système de traitement de la parole de connaissances sur les régularités phonotactiques ou syllabiques de la langue dans les procédures de segmentation lexicale. Les effets obtenus ont été interprétés différemment par les auteurs respectifs de ces études : Vroomen & De Gelder (1999) invoquent la construction d'une représentation syllabique de l'entrée acoustique qui conduirait à une segmentation du signal de parole ne correspondant pas au découpage lexical lorsque le phonème-cible est en position d'attaque syllabique. Cette discordance dans le découpage induirait un délai dans les latences de détection phonémique. McQueen (1998) fait quant à lui appel à la notion de contraintes phonotactiques pour rendre compte d'un allongement des délais d'extraction de mot lorsqu'il n'y a pas alignement entre les frontières phonotactique et lexicale.

Nous souhaitons d'une part mettre en évidence une certaine similarité des composantes phonologiques qui sont supposées avoir été utilisées par les auditeurs dans ces deux études. Pour ce faire, nous approfondirons la distinction qui a été faite entre contraintes phonotactiques et procédures de syllabation. Cette réflexion sur le lien entre contraintes phonotactiques et syllabation permettra, dans le Chapitre 4, de remettre en question les données comportementales obtenues. Nous commençons par décrire les concepts fondamentaux de la phonologie en présentant diverses approches de la structure des représentations phonologiques et des transformations qui leur sont appliquées. Nous présentons alors l'un des concepts centraux, aussi bien dans le cadre de la phonologie que dans celui des travaux sur la perception de la parole : la syllabe. Cette présentation nous permet de mettre au jour un lien étroit entre la conception hiérarchique propre au concept de syllabe et le caractère linéaire des contraintes phonotactiques qui constituent l'une des sources d'information pour la syllabation des séquences de phonèmes. Du fait de cette relation, les travaux précédemment cités semblent se rejoindre sur le plan des processus cognitifs qui sont à l'œuvre dans l'émergence des effets observés. En outre, le caractère linéaire des contraintes phonotactiques nous conduit à envisager la possibilité que des phénomènes corrélatifs aux contraintes phonologiques -mais indépendants de celles-ci- soient impliqués dans les effets comportementaux mis en évidence. Nous concluons sur la nécessité d'adopter une définition opérationnelle du concept de 'légalité' afin d'aborder plus clairement le problème de l'éventuelle intervention de ce que l'on peut considérer comme des 'connaissances phonologiques' dans les processus de segmentation de la parole en mots.

## 1. Représentation sous-jacente et forme de surface

La phonologie est une branche de la linguistique qui se donne pour objectif de comprendre la structure sonore des langues du monde. L'un des postulats des travaux réalisés en phonologie est qu'il existe un format de représentation des sons -la forme phonologique- qui diffère de la forme de surface, c'est à dire la manière dont le message est prononcé -la forme phonétique. Le format de représentation préalable à la réalisation effective des sons est appelé la *représentation sous-jacente*. On peut considérer qu'elle constitue une forme abstraite de la séquence de phonèmes qui va, par l'intermédiaire d'un certain nombre de règles appartenant à ce que l'on nomme une *grammaire*<sup>18</sup>, se transformer pour aboutir à la forme de surface correspondant aux sons effectivement produits. Classiquement, les théories phonologiques proposent deux classes de modèles dans une formulation de l'organisation des sons des langues : elles définissent d'une part la forme des représentations sous-jacentes, c'est à dire le mode de stockage des mots de la langue dans le système phonologique. D'autre part, elles décrivent les opérations à appliquer sur cette représentation pour en dériver une forme phonétique de surface. On a donc affaire à des modèles qui définissent aussi bien des caractéristiques structurales (la forme des représentations) que fonctionnelles (les opérations à appliquer sur ces représentations). Ces différents modèles de la forme des représentations phonologiques et de sa transformation en une forme phonétique de surface sont essentiels à la compréhension du processus inverse : l'appariement entre une forme phonétique de surface et la représentation phonologique correspondante. Les théories développées reposent pour la plupart sur des analyse de corpora. Depuis quelques années, cette approche a par ailleurs été complétée par des études comportementales inspirées des paradigmes de la psychologie cognitive : jeux de langage (Pierrehumbert & Nair, 1995), détection de phonèmes (Cutler, Butterfield, & Williams, 1987), amorçage (Titone & Connine, 1997) qui sont conduites afin de tester les prédictions des modèles développés en phonologie.

### 1.1. Représentations sous-jacentes linéaires vs. hiérarchiques

On peut, en ce qui concerne la forme des représentations phonologiques sous-jacentes, décrire deux classes de modèles qui s'opposent sur le niveau de hiérarchisation de ces représentations. Dans le cadre des travaux initiaux de la phonologie générative proposée par Chomsky & Halle (1968), les représentations phonologiques sous-jacentes ont un format

---

<sup>18</sup> Par analogie avec la grammaire syntaxique.

linéaire. Dans le cadre théorique actuel, on privilégie des représentations multidimensionnelles, non-linéaires.

### 1.1.1. *L'approche initiale de la phonologie générative*

Les travaux originaux de la phonologie générative (Chomsky & Halle, 1968) ont tenté de rendre compte du fonctionnement de la phonologie des langues en ayant recours à des représentations sous-jacentes qui, à tous les niveaux de l'analyse phonologique, présentent une simplicité maximale dans leur structure. Cette simplicité concerne aussi bien l'organisation interne des segments (leur description en termes de traits distinctifs) que leur organisation externe en éléments constitutants de structures hiérarchiques (syllabe, pied, ...) et semble finalement mal adaptée à la description de phénomènes aussi complexes que ceux rencontrés dans les langues humaines. Nous décrivons en premier lieu les choix adoptés pour une représentation linéaire de la structure interne des phonèmes dans ce cadre théorique. Nous passons alors à une présentation du rôle des structures hiérarchiques dans la description de la composante phonologique des langues.

#### 1.1.1.1. Représentation interne des phonèmes

Dans les travaux initiés par Chomsky & Halle (1968) dans *The Sound Pattern of English* (SPE), les segments de la représentation phonologique (les phonèmes) sont décrits comme des séquences de matrices de traits non structurées (cf. Figure 9). Cette absence de structuration induit une indépendance totale entre les traits distinctifs désignant les diverses catégories phonémiques. Le phonème est ainsi envisagé comme une unique matrice de traits, chacune d'elles correspondant à un segment unique de la chaîne de parole (un phonème) et chaque trait pouvant être manipulé indépendamment des autres. Cette conception de la représentation des phonèmes de la chaîne parlée pose des difficultés considérables dans la description des traits

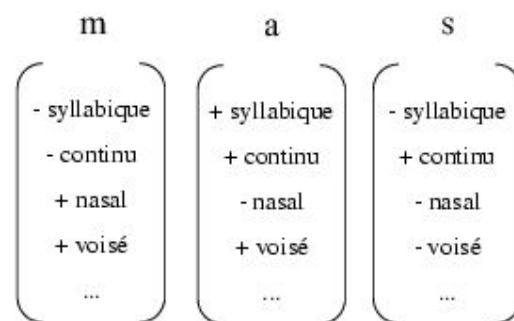


Figure 9 : Représentation interne des segments dans le cadre de SPE (Chomsky & Halle, 1968).

distinctifs. Ainsi, certaines langues utilisent des sons qui présentent une évolution séquentielle de leurs caractéristiques phonémiques. Par exemple, les affriquées se caractérisent par une articulation initiale occlusive (comme /t/) qui se poursuit par une friction (par exemple /ʃ/) mais correspond à un unique segment dans la langue (le phonème /tʃ/). Les mêmes types de modification séquentielle se retrouvent dans les langues à tons : on observe par exemple des voyelles au sein desquelles le ton est ascendant ou descendant. Un ton ascendant peut être décrit comme la succession d'un ton bas et d'un ton haut. Une théorie linéaire des représentations phonologiques dans laquelle les phonèmes seraient envisagés comme des matrices de traits indépendantes et disposées séquentiellement devrait créer de nouveaux traits *ad hoc* pour chaque forme nouvelle observée. Par exemple, Chomsky & Halle (1968) proposent un trait [ $\pm$  delayed release<sup>19</sup>] pour rendre compte de l'existence des affriquées (qui portent la valeur positive [+ delayed release] alors que les occlusives portent la valeur négative du trait). Il est également possible de créer des traits [ $\pm$  rising tone] ou [ $\pm$  falling tone<sup>20</sup>] pour décrire des tons qui évoluent à l'intérieur d'une même voyelle. Nous verrons plus en détails dans la section suivante les modélisations non-linéaires possibles de cette opération en fonction du cadre théorique adopté.

Outre la nécessité de créer des traits *ad hoc* pour rendre compte de certains phénomènes phonologiques, on observe également que le recours à des représentations sous-jacentes linéaires génère des difficultés spécifiques liées à la possibilité qu'ont les segments de présenter des comportements alternatifs. Ainsi, les phonèmes géminés constituent une classe très particulière de phonèmes qui peuvent se comporter aussi bien comme une séquence de deux segments courts que comme une unique consonne longue (Pulleyblank, 1989). Cette alternance dans le comportement du phonème est difficile à modéliser avec la structure des représentations qui est proposée dans SPE (Chomsky & Halle, 1968). En effet, un phonème géminé doit alors être représenté par un ou plusieurs traits distinctifs qui le caractérisent soit en termes de séquence de segments courts, soit en termes d'un segment long. Il semble délicat de décrire une même géminée avec un ensemble de traits distinctifs différents en fonction de son comportement.

#### 1.1.1.2. Représentation de la séquence de phonèmes

Les tenants du courant théorique SPE ont par ailleurs évité de recourir à des structures hiérarchiques comme la syllabe (cf. Goldsmith, 1990, p.103) ou le pied (cf. Pulleyblank, 1989)

---

<sup>19</sup> [ $\pm$  relâchement retardé]

pour rendre compte de l'organisation des phonèmes entre eux. On considère ainsi que c'est une chaîne linéaire d'unités qui subit un ensemble de transformations avant de se constituer en une séquence de segments phonétiques prononçables par le locuteur. Par exemple, si l'on reprend l'exemple donné du mot /medəsẽ/ (médecin), celui-ci subit un ensemble de transformations qui déterminent sa prononciation. Elles pourraient être les suivantes :

$$ə \rightarrow \emptyset / VC \_ CV \quad (1)$$

$$e \rightarrow \varepsilon / \_ C \quad (2)$$

$$[+ \text{voisé}] \rightarrow [- \text{voisé}] / \_ [- \text{voisé}] \quad (3)$$

Ces trois règles<sup>21</sup> conduisent à dériver de la représentation sous-jacente /medəsẽ/ une forme de surface correspondant à la suite phonétique [metsẽ]. Il semble intéressant ici de revenir sur la seconde règle présentée pour rendre compte de la transformation de ce mot. Afin d'expliquer le passage d'un /e/ fermé à un [ε] ouvert, on doit -si l'on souhaite ne pas avoir recours à la notion de syllabe- définir une règle qui soit en mesure d'induire cette transformation à l'aide d'une information purement séquentielle en rendant compte de toutes les exceptions possibles. La description de cette transformation d'un /e/ fermé en [ε] ouvert dans le cadre d'une conception linéaire des représentations phonologiques sous-jacentes se formulerait de la manière suivante dans la règle que nous proposons : un /e/ fermé s'ouvre lorsqu'il précède une consonne. Il est facile de trouver des exceptions à cette formulation. Par exemple, dans le mot /apeti/, la voyelle /e/ est effectivement située entre deux consonnes mais sa forme de surface ne s'ouvre pas. Il est donc nécessaire de modifier la règle (2) afin de préciser les conditions d'application de cette ouverture vocalique. On peut par exemple proposer que la voyelle s'ouvre lorsqu'elle est suivie de deux consonnes :

$$e \rightarrow \varepsilon / \_ CC \quad (4)$$

mais cette règle rencontre encore des exceptions. Le mot /petrir/ se prononce [petrir] alors que le /e/ est suivi de deux consonnes. Il est donc nécessaire d'aller plus loin dans la spécification de la règle en posant que c'est lorsque la voyelle est suivie de deux consonnes obstruantes qu'elle doit s'ouvrir :

<sup>20</sup> [± ton montant] ou [± ton descendant]

<sup>21</sup> Ces règles ne seraient pas nécessairement les mieux adaptées dans le cadre d'un modèle valide de la phonologie du français s'inspirant de la théorie générative. Une quantité considérable de données linguistiques doivent en effet être analysées afin de sélectionner un nombre minimal de règles induisant une quantité minimale d'exceptions.

$$e \rightarrow \varepsilon / \_ [+obstruante] [+obstruante] \quad (5)$$

En réalité, ce processus d'ouverture de la voyelle s'explique très simplement par les caractéristiques de la syllabe qui la porte. En effet, il suffit de dire que la voyelle s'ouvre lorsqu'elle est prononcée dans une syllabe fermée (ou lourde) pour rendre compte de cette transformation. Nous verrons plus en détails les caractéristiques spécifiques de l'entité *syllabe* dans la seconde section de ce chapitre. Néanmoins, avant même d'approfondir la question de ses fonctions et de la description qui en est donnée en fonction des cadres théoriques, on peut déjà définir la syllabe comme une entité typiquement hiérarchique ; selon Goldsmith (1990, p.117) en effet,

[...] syllabation is a process that associates a linear string of segments with a syllable structure<sup>22</sup>.

Le terme de 'structure' syllabique désigne l'organisation de la chaîne de phonèmes en une unité d'ordre supérieur. Il semble ainsi nécessaire, afin de prédire l'organisation phonologique du français avec le moins de règles possibles, d'avoir recours à une indication des frontières d'entités hiérarchiques. La nécessité d'utilisation de la notion de syllabe dans le cadre des règles de la phonologie générative implique donc un relatif abandon du refus d'intégrer des notions de hiérarchisation des séquences de phonèmes dans la grammaire phonologique de la langue. Il reste cependant que la description de l'ouverture du /e/ reste parfaitement descriptible, comme nous l'avons vu, avec des règles ne faisant pas intervenir la syllabe.

Clements & Keyser (1983; cité par Pulleyblank, 1989) décrivent cependant un phénomène qui met en évidence une certaine inadéquation des représentations linéaires pour décrire le fonctionnement des langues. En turc, un certain nombre de règles impliquent un environnement phonémique qui peut présenter deux caractéristiques sans rapport l'une avec l'autre : soit la consonne soumise à la règle phonologique doit être en fin de mot, soit elle doit être suivie d'une autre consonne. Prenons par exemple la règle de dégémiation des consonnes. Si une consonne géminée précède une voyelle, elle n'est pas modifiée. Si cette consonne est suivie d'une consonne ou si elle est prononcée en fin de mot, elle subit par contre une dégémiation. Ces deux conditions constituent, dans le cadre d'une théorie linéaire des représentations phonologiques, deux situations sans rapport l'une avec l'autre. Ces deux environnements, qui ne partagent aucune propriété commune, constituent deux conditions différentes et indépendantes de

<sup>22</sup> La syllabation est un processus d'association d'une chaîne linéaire de segments avec une structure syllabique.



l'application de la règle. Il y a par conséquent disjonction entre deux domaines d'application d'une même règle, l'un phonémique et l'autre lexical. L'existence de disjonctions dans les environnements d'application des règles ne pose pas réellement de problèmes en soi à une théorie linéaire des représentations ; il suffit de créer deux règles distinctes pour rendre compte des deux situations. Cependant, la multitude de situations dans lesquelles Clements & Keyser (1983) observent une disjonction dans l'étude de la phonologie du turc les conduit à remettre en cause la validité d'une approche linéaire.

### 1.1.2. *La phonologie autosegmentale*

De nombreux auteurs remettent ainsi en cause une conception purement linéaire des représentations phonologiques sous-jacentes pour proposer des théories non-linéaires regroupées sous le nom générique de *phonologie autosegmentale* (Goldsmith, 1990; Pulleyblank, 1989). Dans la conception initiée par ce cadre théorique, les représentations phonologiques sous-jacentes sont constituées de plusieurs niveaux de représentation qui portent le nom de 'tire' (de l'anglais *tier*)<sup>23</sup>. La ligne de base, le squelette, constitue une suite de positions segmentales sur lesquelles les éléments des autres lignes peuvent s'ancrer. Les autres tires de la représentation (la tire mélodique, la tire syllabique, la tire tonale, ...) constituent des niveaux de représentations supplémentaires qui rendent compte des différents niveaux d'analyse et des divers problèmes étudiés dans la structure phonologique des langues. La tire mélodique correspond approximativement à la séquence des valeurs phonémiques stockées à l'intérieur de la représentation. La tire syllabique apporte une information concernant la structuration de cette séquence de phonèmes dans la hiérarchie syllabique. La tire tonale (ou accentuelle) représente les informations intonatives attachées à cette séquence. Les phonèmes sont représentés par des arborescences décrivant leur structure en termes de traits distinctifs qui peuvent présenter des dépendances entre eux. Le lien entre un élément et une position du squelette permet de déterminer l'émergence et les modes de manifestation de cet élément dans la forme de surface. Ces positions squelettiques remplissent en effet deux fonctions. Elles permettent aux éléments des autres tires de se manifester lorsqu'ils lui sont liés. Ainsi, seuls les phonèmes ancrés dans le squelette sont exprimés phonétiquement. Elles constituent par ailleurs des unités de temps ou de poids qui, par la structure de leurs connexions avec les autres tires peuvent déterminer par exemple la longueur d'un phonème (Goldsmith, 1990; Pulleyblank, 1989). Cette structuration complexe des représentations phonologiques peut s'appliquer à tout problème de compréhension de la structure sonore des langues. Des problématiques aussi diverses que l'étude des tons des

langues africaines (Rialland & Badjilé, 1989) ou la modélisation de la représentation des traits distinctifs (Clements, 1985) peuvent avantageusement utiliser ce type de modélisation pour rendre compte des données linguistiques.

Ce passage d'un cadre théorique linéaire à des modèles pluridimensionnels trouve sa source dans les difficultés rencontrées par les approches initiales de la phonologie générative pour rendre compte des observations linguistiques. On peut dès lors décrire les représentations phonologiques sous-jacentes comme des structures multidimensionnelles dans lesquelles chaque niveau de description phonologique constitue une facette d'un assemblage d'éléments simples en une structure plus ou moins complexe.

#### 1.1.2.1. Représentation interne des phonèmes

Le recours à des représentations sous-jacentes non-linéaires permet notamment de rendre compte de phénomènes tels que le caractère affriqué des phonèmes ou l'existence de tons alternant dans une voyelle unique sans conduire à l'adjonction de nouveaux traits distinctifs au répertoire. C'est la structure des connexions entre les éléments des diverses tires de la représentation phonologique, mieux que l'invention de nouveaux traits *ad hoc*, qui permet de rendre compte de l'existence de ces structures. Ainsi, une consonne affriquée (cf. Figure 10a) est représentée par une consonne unique dans la ligne mélodique (ce qu'elle est effectivement) qui établit des connexions avec les deux valeurs opposées d'un même trait distinctif ([- continu] pour la partie initiale de l'affriquée et [+ continu] pour sa portion finale). De même (cf. Figure 10b), les tons montants ou descendants s'expriment par la connexion d'un ton bas et d'un ton haut sur une position vocalique ; la voyelle portant le ton montant est par conséquent unique (elle occupe une seule position vocalique) mais elle porte les valeurs de deux tons (un ton haut et un ton bas) se succédant pour générer un contour tonal.

En ce qui concerne les phonèmes géminés, ils se représentent tout simplement par une connexion de l'élément mélodique correspondant au phonème avec deux positions squelettiques

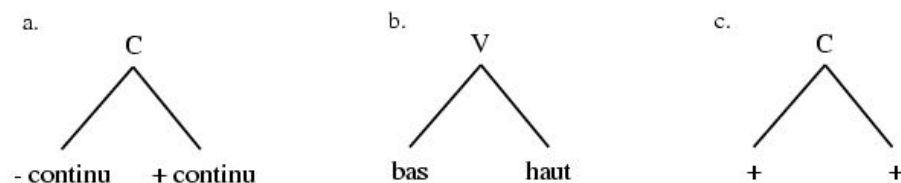


Figure 10 : Représentations non-linéaires décrivant les phénomènes (a.) d'affrication, (b.) de tons montants ou descendants appliqués à un phonème unique ou (c.) de gémination.

<sup>23</sup> On trouve aussi les termes 'ligne' ou 'plan'.

(cf. Figure 10c). Chaque position squelettale représentant une unité segmentale abstraite, il devient inutile d'utiliser des traits spécifiques pour représenter les géminées ; c'est l'ancrage d'un phonème de la ligne mélodique sur deux positions squelettals qui en fait une géminée. Ce mode de représentation facilite par exemple la description du comportement variable des géminées (cf. Section 1.1.1.1 de ce chapitre). En effet, nous avons vu que ces phonèmes peuvent aussi bien adopter le comportement d'une séquence de deux segments courts que celui d'une unique consonne longue. Une représentation interne des phonèmes en termes de matrices de traits nécessite de changer la représentation même du phonème selon l'environnement qui détermine son comportement. Adopter une représentation autosegmentale de la structure des traits distinctifs supprime cette incongruité. Il suffit en effet de dissocier les règles de transformation de la représentation sous-jacente en fonction de la tire sur laquelle elles s'appliquent. Une règle s'appliquant sur la tire mélodique induira un comportement similaire à celui d'une consonne longue alors qu'une règle appliquée sur la tire squelettale déterminera le comportement d'une suite de consonnes courtes. La multidimensionnalité des représentations génère donc une économie considérable dans la description de la structure phonologique des langues. Elle permet en effet de combiner les traits entre eux pour générer des comportements nouveaux. Il est en outre possible d'appliquer les règles de la grammaire sur n'importe quel niveau de la représentation (squelettal, syllabique, tonal, mélodique, ...) et de générer ainsi, avec la même règle et la même représentation, des comportements différents.

Il semble important ici de remarquer que cette problématique de la forme des représentations phonologiques est non seulement essentielle dans le cadre de la phonologie -puisque c'est l'une des questions centrales qui doivent être examinées dans toute théorie de la structure phonologique des langues- mais se trouve intrinsèquement liée aux problèmes soulevés par les études de psychologie cognitive portant sur la perception de la parole. En effet, si l'on suppose l'existence d'un lexique mental unique qui pourrait être utilisé aussi bien dans le cadre de la production que de la perception de la parole et si, par ailleurs, on admet que le mode de stockage des mots dans le système cognitif correspond -au moins en partie- au format proposé dans les modèles élaborés en phonologie, les questions posées influencent directement l'angle sous lequel on abordera l'étude des processus perceptifs.

#### 1.1.2.2. Représentation de la séquence de phonèmes

Ce choix d'une structure sous-jacente multidimensionnelle conduit à des descriptions hautement hiérarchisées de la séquence de phonèmes. Il devient en effet possible de structurer la chaîne de phonèmes à l'aide d'arborescences de type syllabique, moraïque, accentuel ou tonal.

Clements & Keyser (1983) montrent par exemple que le choix de représentations sous-jacentes structurées syllabiquement permet de décrire très simplement la possibilité d'observer un même comportement dans des contextes totalement différents. Si l'on revient sur la règle de dégémination présentée dans la section 1.1.1.2, une formalisation linéaire de la représentation phonologique nécessite de recourir à deux règles. La première consiste à dégémérer la consonne lorsqu'elle est suivie d'une frontière lexicale :

$$[+ \text{gémignée}] \rightarrow [- \text{gémignée}] / \_ \#\# \quad (6)$$

alors que la seconde effectue la même opération lorsqu'elle est suivie d'une consonne :

$$[+ \text{gémignée}] \rightarrow [- \text{gémignée}] / \_ C \quad (7)$$

Dans le modèle proposé par Clements & Keyser (1983), il suffit de recourir à une arborescence syllabique pour rendre compte de cette transformation sans disjonction des domaines d'application. La Figure 11 présente les différents contextes déterminant l'application ou la non application de cette règle de dégémination. La séquence qui subit la transformation est le mot /hakk/ dont la représentation sous-jacente se termine par un /k/ géminé. En (a.), le mot /hakk/ est suivi d'un voyelle ; sa consonne finale conserve donc son caractère géminé. Par contre, en (b.) aussi bien qu'en (c.), cette consonne perd son caractère géminé. Lorsque la consonne est en fin de mot (Figure 11b), c'est à dire lorsque le mot est prononcé isolément, elle est dégémérée en raison de l'impossibilité dans laquelle elle se trouve d'ancrer sa seconde partie dans l'arborescence syllabique<sup>24</sup>. Si elle est suivie d'un mot commençant par une consonne (Figure 11c), elle est confrontée à la même impossibilité et doit également perdre sa gémination. Au contraire, si elle précède une voyelle (Figure 11a), il reste une position d'attaque dans l'arborescence syllabique entre la coda de la première syllabe et le noyau de la seconde syllabe. Il devient alors possible pour cette consonne de s'exprimer phonétiquement puisqu'elle est en mesure de s'ancrer dans la tire syllabique.

Deux ensembles de données favorisent donc le choix de représentations sous-jacentes non-linéaires. Une représentation purement linéaire n'est pas satisfaisante pour décrire la structure interne des phonèmes en termes de traits distinctifs. D'autre part, les règles de transformation de la représentation sous-jacente en une forme phonétique de surface sont grandement simplifiées par la mise en œuvre de structures hiérarchiques comme la syllabe.

<sup>24</sup> Ceci se justifie par le fait qu'une position de coda ne peut porter qu'une seule consonne.

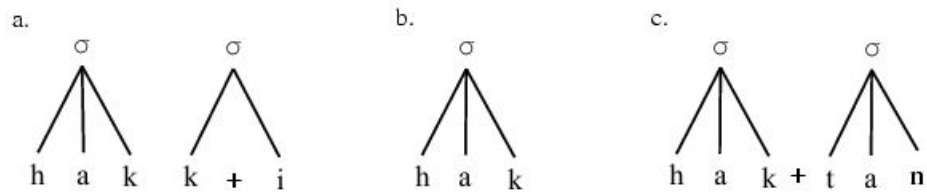


Figure 11 : Représentation de la règle de dégémination du turc dans le cadre de la théorie de Clements & Keyser (1983).

## 1.2. Processus de choix d'une forme de surface

Outre la modélisation structurale, dont l'objet est de rendre compte de la forme des représentations sous-jacentes, il est nécessaire de décrire le fonctionnement du système qui permet de passer de ces représentations sous-jacentes à une forme phonétique de surface. Or cette conception d'une opération consistant à dériver une représentation d'une autre est actuellement remise en question par les tenants de la Théorie de l'Optimalité.

### 1.2.1. Transformation d'une représentation sous-jacente

Comme nous l'avons vu depuis le début de ce chapitre, on suppose en phonologie l'existence de deux types de représentations. La première, la représentation sous-jacente, est une entité sur laquelle vont s'appliquer un ensemble de règles de transformation. On peut considérer la somme des représentations sous-jacentes de la langue comme un système qui serait stocké dans le système langagier des locuteurs (par exemple dans le lexique). Le second type de représentations, la forme phonétique de surface, est dérivé des représentations sous-jacentes par l'application des règles de la grammaire phonologique. Cette forme phonétique de surface serait dérivée de ces règles et ne resterait pas stockée par le système phonologique. C'est donc une transformation de la représentation sous-jacente, transformation réalisée par le biais des règles de la phonologie, qui permet d'aboutir à la forme phonétique de surface.

### 1.2.2. Sélection de la sortie en fonction des caractéristiques souhaitées

Dans le cadre de la Théorie de l'Optimalité, le processus de constitution de la forme phonétique de surface n'est pas une opération de dérivation par règles. Contrairement à la position selon laquelle on transforme une représentation sous-jacente en une forme de surface, la Théorie de l'Optimalité (OT pour Optimality Theory, Prince & Smolensky, 1993), propose que l'on sélectionne la forme de surface parmi l'ensemble des formes de surface possibles. Dans les descriptions de la composante phonologique que nous avons données au préalable, les

représentations sous-jacentes étaient soumises à des règles de transformation qui permettaient de parvenir progressivement à une forme correspondant à la sortie phonétique. Dans le cadre de la Théorie de l'Optimalité, on ne parle plus de règles mais de contraintes. Ces contraintes ne servent pas à modifier la structure de la représentation sous-jacente mais à sélectionner le candidat optimal pour la sortie. La composante phonologique ne modifie donc pas ici la représentation phonologique sous-jacente mais choisit un candidat de surface en sélectionnant, parmi l'ensemble des candidats possibles, celui qui respecte au maximum l'ensemble des contraintes définies dans la grammaire. On considère actuellement que les contraintes décrites sont universelles, c'est à dire qu'elles existent dans toutes les langues du monde. Seul leur ordre de priorité permet de différencier les langues entre elles. Ainsi, un changement dans l'ordre des contraintes permet de modifier l'inventaire des formes phonétiques observées dans la langue. Deux classes de contraintes peuvent être proposées : les contraintes *phonotactiques* déterminent une propriété à laquelle doit se conformer la sortie. Dans la seconde catégorie de contraintes, que l'on appelle en anglais *faithfulness constraints*, on considère que la sortie, la forme phonétique de surface, doit être le plus proche possible de l'entrée, la représentation sous-jacente, pour une propriété déterminée.

Après cette rapide introduction aux diverses problématiques rencontrées en phonologie concernant la description interne des segments, leur possible organisation en constituants de structures hiérarchiques et les modalités permettant d'aboutir à une forme phonétique de surface, nous revenons plus longuement sur une structure hiérarchique qui nous intéresse particulièrement dans le cadre de cette thèse : la syllabe.

## 2. La syllabe

Nous avons entrevu, dans la section précédente, l'intérêt que peut représenter la syllabe pour décrire -dans le cadre d'une théorie générative non-linéaire- la structure des représentations phonologiques sous-jacentes et faciliter la description de règles de transformation de ces représentations en formes de surface. Nous verrons que ce concept est également utilisé dans la Théorie de l'Optimalité. Il reste cependant que, malgré son utilité, la syllabe est un concept mal défini qui, même si l'on y recourt fréquemment, n'en reste pas moins flou (Arnold, 1956). Ce problème s'illustre notamment par l'observation suivante : tout locuteur natif est capable de compter avec précision le nombre de syllabes d'une séquence de parole prononcée dans sa langue... mais il est souvent bien en peine d'indiquer avec certitude la localisation des frontières syllabiques. Ceci semble mettre en évidence l'un des problèmes liés à cette notion de syllabe : on

peut facilement pressentir ce concept comme un indicateur de la structuration de la chaîne de parole (Arnold, 1956 utilise l'expression anglaise 'dealing with syllabics' pour désigner le type de connaissances nécessaires à la mise en œuvre de cette capacité) ; mais son analyse en termes de constituants -ainsi que le découpage syllabique lié à l'identification précise des constituants ('dealing with syllables' selon Arnold, 1956)- induisent le recours à des connaissances tout à fait différentes. C'est cette dernière composante, décrire des règles permettant de localiser avec certitude les frontières syllabiques, qui pose le plus problème dans le cadre de la description phonologique de la syllabe.

## 2.1. Description de la notion de syllabe

La description de l'entité correspondant à ce que l'on appelle la syllabe nécessite d'adopter plusieurs points de vue. En premier lieu, nous verrons à quoi sert la syllabe dans la description de la composante phonologique d'une langue. Nous illustrerons ce rôle par des exemples mettant en évidence la multiplicité des langues dans lesquelles cette notion intervient et la non-adéquation d'une description en termes de séquences purement linéaires de segments. Nous présenterons ensuite diverses conceptions de cette entité. Nous aborderons alors le problème du découpage de la chaîne de segments en syllabes en mettant en évidence les difficultés que pose la localisation des frontières syllabiques. Cette étape de l'analyse nous conduira à poser la question de la validité d'une conception hautement hiérarchique de la structuration syllabique.

### 2.1.1. *A quoi sert la syllabe ?*

Cette question nous semble importante dans la mesure où l'approche initiée dans SPE (Chomsky & Halle, 1968) consiste à se dispenser de recourir à des entités hiérarchiques pour décrire la structure phonologique des langues. Il faut donc, avant de poser la question de ce qu'est la syllabe, être raisonnablement certain de son intérêt dans la théorie phonologique. Nous avons vu dans les sections 1.1.1.2 et 1.1.2.2 quelques illustrations mettant en évidence un rôle possible de la structure syllabique dans la description des règles de transformation de la représentation sous-jacente. Dans le premier exemple que nous avons donné concernant l'intérêt de la syllabe pour la description des règles phonologiques, nous avons choisi la règle d'ouverture du /e/ dans une syllabe fermée. Nous avons cependant admis qu'une description adéquate en termes d'une séquence linéaire de segments pouvait parfaitement rendre compte de cette opération. Les travaux de Clements & Keyser (1983) illustrent néanmoins une incongruité relative à la disjonction des environnements d'application des règles du turc dans le cadre d'une théorie qui ferait abstraction de la syllabe.

Si la règle d'ouverture du /e/ en français peut tout à fait s'expliquer à l'aide d'une séquence linéaire de segments et que les règles du turc présentées par Clements & Keyser (1983) peuvent se satisfaire -même si cela semble incongru- d'une disjonction des environnements, un nombre considérable d'autres langues ont été décrites dans lesquelles la syllabe apparaît de nouveau comme une référence essentielle, voire incontournable, de la description phonologique. Dans les langues, la syllabe peut remplir deux fonctions (Goldsmith, 1990). Elle sert d'une part à réguler les règles d'association segmentales, c'est à dire à influencer les conditions d'application des règles de transformation de la représentation sous-jacente. Cette régulation peut s'opérer en fonction de la structure de la syllabe à laquelle appartient le segment qui est soumis à la règle. Ainsi en est-il de la règle d'ouverture du /e/ dont nous avons parlé précédemment pour le français. La même catégorie de phénomènes s'observe en français québécois. Dans cette langue, il existe une règle de relâchement du /i/ dans une syllabe fermée. Par exemple, le /i/ de 'petit' reste tendu dans sa forme masculine ([pti]) mais se relâche dans sa forme féminine qui place le /i/ sous-jacent dans une syllabe fermée ([ptit]). La syllabe permet également de réguler les transformations phonologiques en fonction de la place du segment dans la syllabe. En allemand, toute obstruante perd son caractère voisé lorsqu'elle occupe une position de coda syllabique<sup>25</sup>. Alors que la forme féminine *freundin* -'amie'- se prononce [frɔjndən], le /d/ sous-jacent se dévoise dans la forme masculine *freund* -'ami'- dans laquelle il occupe la position syllabique de coda ; *freund* se prononçant [frɔjnt]. Cette catégorie de règles peut difficilement être décrite en faisant abstraction d'une structuration hiérarchique de la séquence de segments.

La syllabe permet d'autre part de déterminer la position lexicale de l'accent. En Yup'ik, une langue parlée dans certaines régions des USA et de l'ancienne URSS, l'accent est porté par la première syllabe fermée. Si la première syllabe d'un mot est ouverte, c'est la syllabe suivante qui porte l'accent. Ce phénomène met en évidence un lien étroit entre les diverses composantes phonologiques des langues. Ainsi, la composante correspondant à la structuration syllabique de la chaîne de segments influence la composante accentuelle. Dans le cadre de la phonologie métrique (*metrical phonology*) par exemple, la syllabe constitue un concept essentiel pour décrire les alternances rythmiques.

---

<sup>25</sup> coda = fin de syllabe (cf. section 2.1.2.1 pour une description des constituants syllabiques)



### 2.1.2. Qu'est-ce que la syllabe ?

Puisque cette notion de syllabe constitue un élément incontournable de la description des langues, tout recours à ce concept nécessite qu'il soit défini au préalable. Or définir la syllabe est une tâche laborieuse. Le phonème est une notion *relativement* simple à définir dans le cadre de la linguistique étant donné qu'il constitue la plus petite unité distinctive de la langue. Les séquences /ba/ et /ga/ constituant deux formes phonétiques qui s'opposent sur une certaine dimension pour les locuteurs du français, la paire /b/-/g/ présente un statut phonémique en français. Le phonème peut en outre se définir comme le plus petit segment interchangeable dans la langue du fait de l'opposition possible entre /bi/ et /ba/ qui implique par exemple de distinguer les parties initiale et finale de /ba/. Le phonème présente par conséquent une certaine réalité dans la description de la langue puisqu'il en est le fondement même ; décrire une langue consiste à l'analyser en termes de distribution des unités segmentales minimales. La syllabe au contraire n'a pas de statut linguistique clair. On considère classiquement que cette entité constitue une *unité phonologique* à part entière dans le sens où elle joue un rôle dans la description de la langue. Elle ne peut cependant pas être décrite comme une unité segmentale distinctive puisqu'elle en serait elle-même constituée, pas plus qu'il n'est possible d'en donner une description physique en termes de frontières acoustiques ou de pauses articulatoires. Lorsqu'ils tentent de définir le concept de 'syllabe', les auteurs décrivent pour la plupart les règles permettant de découper une chaîne de segments en groupes syllabiques ou présentent leur propre conception de la structure syllabique interne. Par exemple, Hooper (1972) introduit son étude sur le rôle de la syllabe dans la théorie phonologique par ces termes :

The purpose of this paper is [...] to suggest a universal definition [of the syllable] in terms of conventions for the placement of syllable boundaries<sup>26</sup>.

Une définition qui nous semble importante est cependant donnée par Fujimura (1990) :

I think a syllable should be defined as a minimal unit that is utterable in isolation at the phonetic level, and any use of the term should be in some way consistently related to this phonetic notion<sup>27</sup>.

<sup>26</sup> L'objet de cet article est [...] de suggérer une définition universelle [de la syllabe] en termes de conventions concernant la localisation des frontières syllabiques.

<sup>27</sup> Je pense que la syllabe devrait être définie comme l'unité minimale d'articulation au niveau phonétique, et tout usage de ce terme devrait d'une manière ou d'une autre être intimement rattaché à cette notion phonétique.

Nous avons remarqué précédemment qu'aucune pause articulatoire ne permet de localiser des frontières syllabiques dans un signal de parole continue. Il apparaît cependant que, si l'on concentre notre intérêt sur ce qu'est une syllabe prononcée isolément, la conception avancée par Fujimura (1990) fournit des renseignements essentiels. La syllabe constituerait selon cette proposition une unité minimale d'articulation. Cette image nous semble aisément assimilable pour comprendre que cette entité 'syllabe' est centrée sur quelque chose qui peut être articulé en isolation (par exemple une voyelle) et peut se composer, autour de cet élément articulable, d'éléments qui ne sont pas nécessairement prononçables isolément mais qui, adjoints au précédent, le deviennent. Ceci nous paraît, en élidant la question du statut phonétique ou sous-jacent de l'entité syllabe, un préliminaire correct à une description de la syllabe. Nous n'irons pas plus loin dans cette tentative de définition. Il nous semble plus naturel, du fait de la difficulté intrinsèque à cette tâche, de présenter les diverses propositions qui ont été avancées concernant sa description afin d'appréhender cette notion à partir de la représentation qui en est donnée dans les diverses approches théoriques pour lesquelles c'est l'un des concepts centraux. Nous commençons par présenter deux des conceptions essentielles qui se sont classiquement opposées pour ce qui a trait à ce qu'est une syllabe : l'approche en termes de constituants (*constituent approach*) et celle en termes de courbes (*curve approach*). Nous terminons par la présentation de la théorie moraique qui, dans une certaine mesure, bouleverse la vision hautement hiérarchisée à laquelle nous sommes parvenus depuis quelques années.

#### 2.1.2.1. Les constituants

On peut décrire la syllabe comme une structure hiérarchique dont les phonèmes sont les *constituants*. Ces éléments simples se regroupent en une unité d'ordre supérieur et y occupent des positions qui possèdent chacune un statut spécifique. Cette spécificité du statut des différentes positions syllabiques conduit à attribuer à chaque élément de l'arborescence une

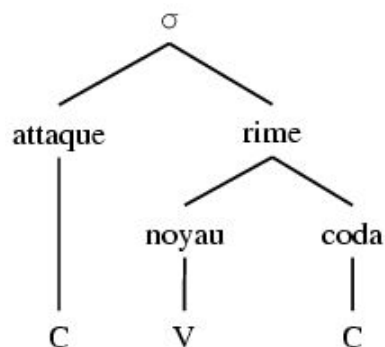


Figure 12 : Représentation hautement hiérarchique de la syllabe en constituants phonémiques remplissant les fonctions d'attaque, noyau et coda.

fonction différente. La représentation syllabique d'un mot adoptera la forme représentée dans la Figure 12. Les constituants syllabiques sont, dans le cadre de la conception la plus répandue de la syllabe (Fudge, 1987) : l'attaque (*onset*, consonne ou groupe de consonnes initial) et la rime (*rhyme*, les phonèmes restants), cette dernière se dissociant en noyau (*nucleus* ou *peak*, généralement la voyelle) et coda (la consonne finale). Cette approche en termes de constituants représente la syllabe comme un groupe de segments phonémiques agencés à l'intérieur d'une structure hautement hiérarchisée et privilégie par conséquent l'aspect interne de cette unité.

#### 2.1.2.2. Les courbes syllabiques

L'une des approches alternatives consiste à définir la syllabe comme une alternance d'accroissement et de diminution d'une variable quelconque de la parole. Par exemple, on peut considérer que la succession des syllabes dans l'activité de parole correspond à une alternance de sons caractérisés par une quantité importante d'énergie (niveau d'intensité acoustique élevé, ce qui est par exemple le cas des voyelles) et de faible énergie (intensité acoustique restreinte comme dans les consonnes). Ces courbes peuvent également se définir en termes articulatoires à partir de la pression engendrée par la constriction (faible pression pour les voyelles et haute pression pour les occlusives). Cette conception de la syllabe se fonde sur une analyse de ses propriétés externes (notamment acoustiques ou articulatoires) plus que sur le statut des éléments la constituant. Elle conduit cependant à l'émergence de difficultés concernant le choix d'un indice acoustique ou articulatoire pertinent (cf. Section 2.2.1.2).

#### 2.1.2.3. La syllabe moraique

Récemment, un certain nombre d'auteurs (cf. notamment Hyman, 1985; Hayes, 1989) ont proposé, dans le cadre d'une approche en constituants, une conception alternative de la syllabe qui consiste à réduire le niveau de hiérarchisation de l'arborescence tout en conservant une conception non-linéaire de la représentation sous-jacente. On parle, pour désigner cette

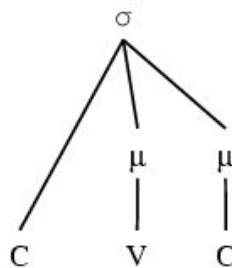


Figure 13 : Représentation moraique de la syllabe.

proposition, de *structures syllabiques plates*<sup>28</sup> (Pierrehumbert & Nair, 1995). Cette position, connue sous le nom de *théorie moraïque*, consiste à relier directement l'attaque de la syllabe au nœud syllabe. Les autres phonèmes la constituant sont quant à eux liés aux nœuds moraïques : les mores. Il n'existe pas, dans ce modèle de la syllabe, de constituants présentant un niveau de hiérarchisation plus complexe. Dans le cadre de la théorie moraïque, la more constitue une unité de poids. Elle est donc bien plus qu'un constituant se substituant au noyau ou à la coda.

Cette unité permet par exemple de rendre compte très efficacement du comportement des voyelles longues ou de la similarité de comportement entre des syllabes comportant une consonne en position de coda (CVC) et des syllabes portant une voyelle longue (CV:). Dans la Figure 14 on observe que, pour les mêmes mots, les représentations dérivées de la théorie moraïque sont plus simples que celles exprimées en termes de constituants. Ainsi, les deux consonnes initiales du mot français /trak/ n'ont pas besoin de brancher sous le nœud d'attaque et peuvent être directement reliées au nœud syllabique. De même, la représentation d'une voyelle longue comme dans le mot anglais /trif:/ (*tree*, arbre) se passe de branchements doubles (de la voyelle vers le squelette et de ces positions squelettales vers le nœud noyau). Il suffit de faire

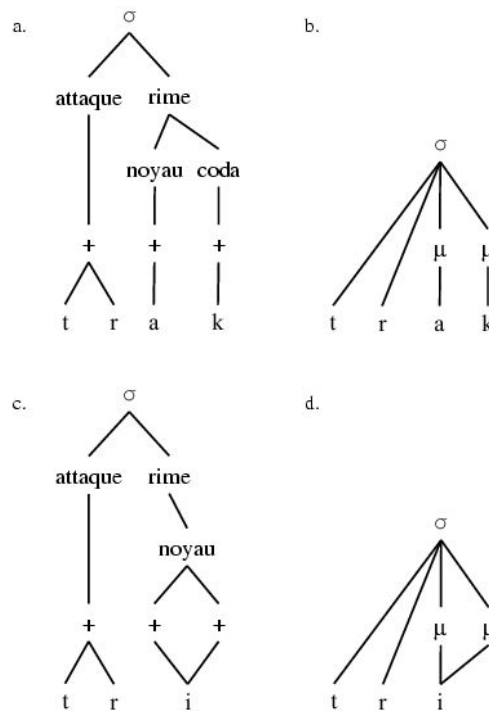


Figure 14 : Comparaison de la représentation de deux mots (l'un français, 'traque' ; l'autre anglais, 'tree') en fonction du choix de la structure syllabique (arborescence hautement hiérarchique en a. et c., syllabe moraïque en b. et d.).

<sup>28</sup> *flatter structures for the syllable*

brancher la voyelle vers les deux positions moraiques (notées  $\mu$ ) pour rendre compte à la fois de son allongement et de sa place dans la syllabe. En effet, une voyelle longue n'occupe pas la position codique mais celle de noyau ; il n'est donc pas envisageable de faire seulement brancher la voyelle vers les positions squelettales et de conserver les connexions du squelette vers les positions de noyau et de coda. Dans la théorie moraique, poids et position syllabique sont donc intrinsèquement liés. Dans ce cadre là, une syllabe portant une voyelle longue (CV:) se comporte exactement comme une syllabe fermée (CVC). On parle alors, pour mettre en évidence leur similarité, de syllabe 'lourde'. Une syllabe lourde comporte deux mores ; une syllabe légère n'en comporte qu'une. Par ailleurs, le fait qu'une consonne ou un groupe de consonnes en position d'attaque puisse présenter un comportement différent de celui manifesté par la même consonne ou suite de consonnes en coda s'explique par la source du branchement qui provient directement du nœud syllabique en attaque alors qu'il trouve sa source dans les positions moraiques en coda.

## 2.2. Déterminants de la structuration syllabique

Après avoir présenté les diverses conceptions de la structure interne de la syllabe dans le but de fournir les fondements d'une compréhension de ce que peut être cette entité couramment utilisée en phonologie, il est important de distinguer ce qu'est la syllabe des opérations permettant de localiser ses frontières. Que l'on adopte une approche en constituants aussi bien qu'une approche de courbe syllabique, il faut -afin de segmenter une séquence de phonèmes en syllabes- localiser les frontières qui les séparent. Le processus de syllabation par lequel (1) les segments constitutifs de l'arborescence syllabique acquièrent une position spécifique ou (2) les phonèmes de la chaîne parlée se distribuent en fin ou en début de syllabe en fonction de leurs caractéristiques acoustiques ou articulatoires est donc une composante essentielle du problème qui nous concerne, c'est à dire la capacité qu'auraient les locuteurs natifs de recourir à des connaissances sur la structuration phonologique de la chaîne parlée dans les processus de segmentation de la parole en mots.

### 2.2.1. *Principes d'organisation indépendants des caractéristiques des phonèmes*

Plusieurs approches ont été proposées afin de dériver une séquence syllabiquement structurée à partir d'une chaîne linéaire de phonèmes. La première catégorie de propositions adopte l'approche classique des dérivations par règles, celles-ci reposant sur des principes indépendants des caractéristiques intrinsèques des phonèmes (principes de l'attaque obligatoire, de l'attaque maximale). Des approches inspirées de la théorie de l'optimalité envisagent au

contraire que le processus de syllabation repose sur la satisfaction de contraintes ordonnées permettant de sélectionner la forme de sortie parmi l'ensemble des candidats possibles. Finalement, un certain nombre d'influences supplémentaires doivent être prises en compte afin de refléter adéquatement les procédures de syllabation de la chaîne de phonèmes : des influences liées à la morphologie, à l'accent (durées vocaliques notamment) et à l'intervention de contraintes phonotactiques.

#### 2.2.1.1. Approche par règles

L'une des approches concernant le mécanisme de structuration syllabique d'une chaîne phonémique consiste à proposer des règles permettant de transformer une forme d'entrée composée d'une chaîne linéaire de segments en une séquence de segments qui acquerraient un statut spécifique dans la hiérarchie syllabique.

##### 2.2.1.1.1. Principe de l'Attaque Obligatoire

Hooper (1972) propose un principe de syllabation selon lequel toute syllabe doit, tant que cette procédure ne viole pas les contraintes phonotactiques de la langue, comporter une attaque. Dans une séquence CVCVC (par exemple /pətit/), ce principe conduit à une syllabation de la suite de phonèmes en CV#CVC car, toute syllabe devant comporter une attaque, la consonne médiane se situe nécessairement à l'attaque de la seconde syllabe. Cette opération ne s'applique cependant pas si la suite médiane CV constitue une séquence phonotactiquement illégale dans la langue. Pour des séquences contenant plus d'une consonne médiane, ce principe n'est pas suffisant. Ainsi dans une suite  $CVC_1C_2VC$ ,  $C_2$  doit nécessairement appartenir à la seconde syllabe, lui permettant de disposer d'une attaque. Par contre, le statut de  $C_1$  est indéterminé. Elle peut tout aussi bien constituer la coda de la première syllabe que se regrouper avec  $C_2$  à l'attaque de la seconde syllabe. Il est par ailleurs impossible de choisir une solution définitive *a priori* qui consisterait à placer  $C_1$ , quelle qu'elle soit, en coda (ou l'inverse) puisqu'en français certaines séquences CVCCVC se syllabent en CVC # CVC (par exemple, 'taxi', /tak # si/) alors que d'autres présentent une structure CV # CCVC (par exemple, 'sacré', /sa # kre/). Des règles supplémentaires sont donc nécessaires pour améliorer la validité de ce principe.

##### 2.2.1.1.2. Principe de l'Attaque Maximale

L'objet du Principe de l'Attaque Maximale (Selkirk, 1982; cf. Goldsmith, 1990 pour une revue) est de fournir les fondements d'une prédiction plus complète des structures syllabiques observées en fonction de la séquence linéaire de phonèmes considérée. Conformément au *Maximum Onset Principle*, l'on doit insérer un nombre maximal de consonnes en position

d'attaque syllabique, ceci à condition que cette opération ne viole pas les contraintes phonotactiques de la langue. Dans une séquence CVCCCVC (par exemple, /syrpriz/), il devient possible de décider de la localisation adéquate de la frontière en fonction de la suite de consonnes médiane. L'*Obligatory Onset Principle* ne fournit pas d'information quant au nombre de consonnes à placer en attaque de syllabe, il n'est donc pas possible de choisir entre les diverses configurations possibles. Le *Maximum Onset Principle*, au contraire, contraint un maximum de consonnes à occuper cette position. Ainsi, dans la suite CVCCCVC, les trois consonnes médianes seront positionnées à l'attaque de la seconde syllabe pour autant que les contraintes phonotactiques de la langue le permettent. Si, dans une séquence  $CVC_1C_2C_3VC$  provenant d'une langue donnée, les suites  $C_1C_2$ ,  $C_2C_3$  et  $C_3V$  sont phonotactiquement légales, alors la séquence de phonèmes sera découpée en  $CV \# C_1C_2C_3VC$ . Par contre, si  $C_2C_3$  constitue un groupe de consonnes illégal dans cette langue, la suite de segments se découpera en  $CVC_1C_2 \# C_3VC$ . Le *Maximum Onset Principle* fournit donc une méthode de découpage syllabique plus contrainte que l'*Obligatory Onset Principle*. Elle ne permet cependant pas de prédire intégralement la structuration syllabique de l'ensemble des mots possibles.

#### 2.2.1.2. Optimalité des formes syllabiques

Le développement de la Théorie de l'Optimalité a donné lieu à des propositions spécifiques concernant les procédures de syllabation. Dans ce cadre théorique, on considère que le processus de structuration syllabique correspond à un mécanisme de satisfaction de multiples contraintes qui sont ordonnées selon des critères spécifiques à chaque langue, cette hiérarchie des contraintes permettant de donner lieu à des structures syllabiques différentes en fonction des langues. Tesar (1995) propose un algorithme de syllabation reposant sur 5 contraintes :

ONS	Les syllabes doivent comporter une attaque
NOCODA	Les syllabes ne doivent pas comporter de coda
PARSE	Les segments doivent être rattachés à des positions syllabiques
FILL <sup>Nuc</sup>	Les positions nucléiques doivent être remplies (avec une voyelle)
FILL <sup>ONS</sup>	Les positions d'attaque doivent être remplies (avec une consonne)

Ces contraintes peuvent être transgressées. La sortie doit seulement être optimale, c'est à dire respecter au maximum l'ensemble de ces contraintes (et notamment celles qui sont le plus haut placées) mais elle peut en transgresser quelques unes. En fonction des langues, leur ordre peut varier, ce qui détermine quelles sont les syllabes optimales dans une langue donnée.

Quel que soit le cadre théorique adopté, on notera qu'une approche cherchant à rendre compte des procédures de syllabation par des méthodes indépendantes des phonèmes en jeu fait des différentes positions syllabiques que sont l'attaque, le noyau et la coda des concepts centraux de son analyse ; induisant de fait une conception hiérarchique de la syllabe. D'autres modèles, dérivés de l'approche en termes de courbes, proposent de faire reposer la procédure de syllabation sur des caractéristiques intrinsèques aux phonèmes. Ces conceptions reposent moins sur le statut hiérarchique de la syllabe que sur sa structuration comme unité phonologique. Elles se situent par ailleurs à l'interface de la phonologie et de la phonétique puisqu'elles insistent sur l'importance des caractéristiques propres des phonèmes dans leur organisation en syllabes.

### 2.2.2. Principes reposant sur les caractéristiques des phonèmes

Que l'on se situe dans un cadre classique de dérivation par règles (OOP, MOP) ou de satisfaction de contraintes (OT), les mécanismes proposés ne prennent pas en considération les caractéristiques intrinsèques des phonèmes. L'objectif est en effet de prédire la structure syllabique d'une séquence indépendamment de ces caractéristiques. Certaines approches proposent cependant de recourir explicitement à ces caractéristiques afin de prédire la structure syllabique d'une séquence phonémique.

#### 2.2.2.1. Echelle de sonorité

Les deux principes précédents reposent sur des règles d'association syllabique évitant de recourir à des caractéristiques spécifiques des phonèmes considérés. Les procédures de syllabation fondées sur un principe de sonorité (cf. notamment l'approche en termes de courbes) reposent au contraire explicitement sur une description des phonèmes en termes de *sonorité* afin de prédire la structuration syllabique d'une chaîne phonémique (cf. Clements, 1990 ; Klein, 1993). Une syllabe *bien formée* correspondrait à une suite de phonèmes présentant un accroissement puis une diminution *monotones* de la sonorité des segments qui la constituent. On peut définir la sonorité aussi bien selon une approche acoustique qu'articulatoire. Une définition acoustique de la sonorité pourrait ainsi prendre en considération la quantité d'énergie mesurable dans le segment acoustique correspondant approximativement au phonème. Par exemple, une voyelle se caractérise par une enveloppe d'intensité contenant beaucoup d'énergie (cf. Figure 15a) alors qu'une consonne se manifeste par une quantité d'énergie relativement faible (cf. Figure 15b). Des mesures articulatoires pourraient également être mises en œuvre afin d'estimer la sonorité d'un segment : ont par exemple été proposés les indices d'ouverture de l'appareil vocal ou de taille de la chambre de résonance à travers laquelle passe le flux d'air qui donne lieu au segment considéré. Il a également été proposé une échelle de consonanticité qui est



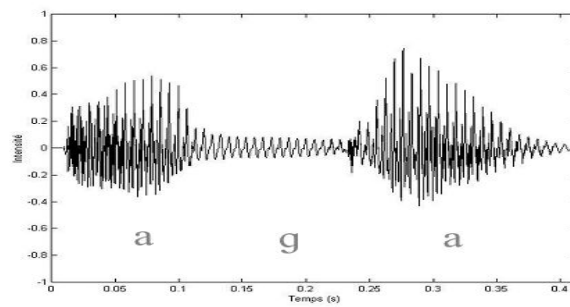


Figure 15 : Enveloppe d'intensité du logatome /aba/ mettant en évidence la différence de quantité d'énergie caractérisant deux classes de phonèmes : les occlusives sonores (comme /g/) et les voyelles.

exactement l'inverse de celle de sonorité (Klein, 1993). Cet indice pourrait quant à lui être dérivé d'une représentation spectrographique du signal reposant sur les caractéristiques du spectre de chaque segment.

A partir d'une comparaison des structures syllabiques observées dans diverses langues, plusieurs échelles de sonorité (ou de consonanticité) ont été proposées qui devraient permettre de prédire la structuration syllabique d'une chaîne linéaire de segments. Deux exemples en sont donnés dans la Figure 16. On observe, de haut en bas, une diminution de la sonorité ou un accroissement de la consonanticité en fonction des types de phonèmes. On pourrait, selon cette

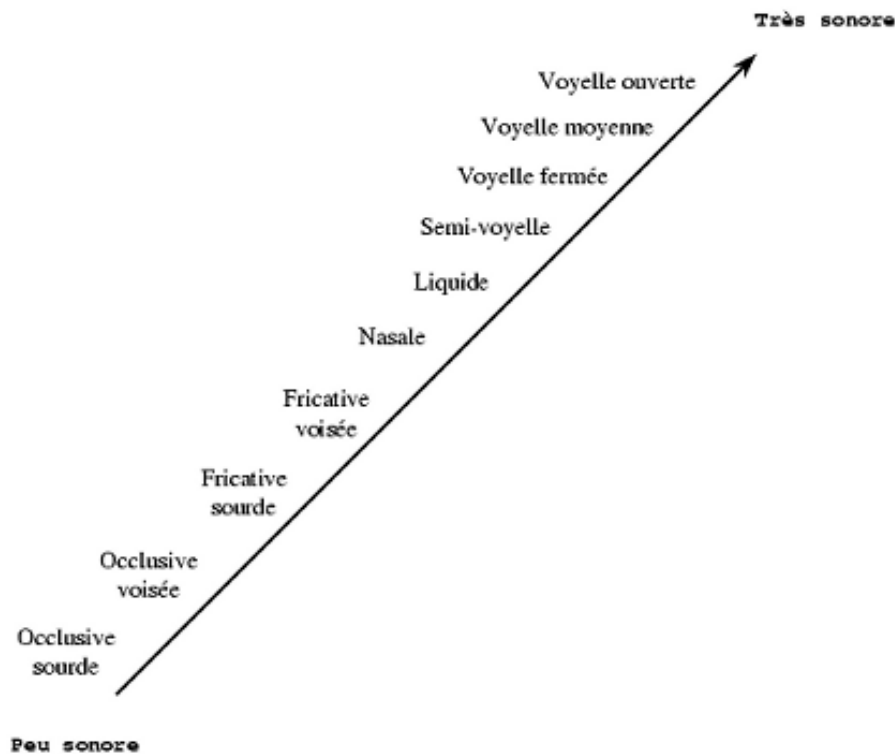


Figure 16 : Distribution de diverses classes de phonèmes selon l'échelle de sonorité.

échelle, déterminer directement à partir des caractéristiques intrinsèques des segments la structure syllabique d'une chaîne de phonèmes. Sur la base d'une échelle donnée de sonorité, la procédure de syllabation consiste à organiser la séquence de segments de telle sorte que le profil de sonorité croisse de façon monotone jusqu'au sommet syllabique puis décroisse alors de manière monotone jusqu'à la coda. Ce principe repose, de même que l'OOP ou le MOP, sur la condition de légalité phonotactique des séquences de phonèmes intégrées dans la syllabe.

En réalité, il apparaît que cette échelle de sonorité ne trouve aucune justification dans les caractéristiques acoustiques ou articulatoires des phonèmes (Clements, 1990). Il ne semble pas possible en effet de déterminer la sonorité d'un phonème à partir d'une analyse de ses caractéristiques physiques. Cette échelle hypothétique n'a été constituée qu'*a posteriori*, elle est déduite de la description des structures syllabiques rencontrées dans plusieurs langues mais, à l'exception de certaines distinctions majeures (par exemple la différence entre occlusive et voyelle), la plupart des classes de phonèmes sont difficilement classifiables à partir de mesures physiques de la sonorité. Par ailleurs, certains phonèmes très semblables sur le plan de leur articulation et de leur niveau dans l'échelle de sonorité présentent des comportements totalement différents dans la structuration syllabique. En français par exemple, les occlusives coronales /t/ et /d/ se comportent tout à fait différemment des autres occlusives bien que l'échelle de sonorité leur attribue un statut équivalent. Par exemple, /t/ ne peut pas être regroupé à l'attaque syllabique avec /l/ alors que l'attaque /gl/ est quant à elle bien formée. Il reste cependant que cette échelle de sonorité permet de modéliser la syllabation d'une chaîne de parole avec plus de précision que les principes de l'Attaque Obligatoire ou Maximale. Clements (1990) en conclut que cette échelle de sonorité, déterminée *a posteriori* et ne correspondant à aucune réalité physique établie, correspondrait non à une variable mesurable avec précision dans le signal (ou au cours de la production) mais à des caractéristiques sous-jacentes des phonèmes, caractéristiques abstraites relevant de la phonologie mais pas de la phonétique. Cette dernière composante fournit cependant les fondements de distinctions primaires entre quelques classes phonémiques (obstruante, liquide et voyelle par exemple).

#### 2.2.2.2. Similarité

Indépendamment des travaux sur la syllabation, Frisch (Frisch, Broe, & Pierrehumbert, soumis; Frisch, 1996) propose d'avoir recours à une mesure de similarité entre phonèmes pour rendre compte des contraintes phonotactiques régulant la co-occurrence des consonnes. Plus deux phonèmes seraient similaires, moins ils auraient tendance à être juxtaposés dans les mots de la langue. Ce taux de similarité pourrait être évalué à partir de la description des phonèmes en

termes de traits distinctifs. En outre, les auteurs proposent que cet indice de similarité serait structuré, c'est à dire que certains traits pourraient déterminer le partage d'autres traits. En arabe, langue à morphologie concaténative, la racine verbale est uniquement composée de consonnes. L'obtention d'une forme verbale déterminée passe par l'insertion, au cœur de cette racine verbale consonantique, de voyelles jouant le rôle d'affixe (cf. Figure 17). Une contrainte nommée *OCP-Place* (McCarthy, 1988) empêche de combiner des consonnes homorganiques<sup>29</sup> côte à côte à l'intérieur de la racine verbale, donc de la tire consonantique. Frisch et al. (soumis) observent que cette contrainte n'est pas binaire : elle s'applique avec plus ou moins d'intensité en fonction du taux de similarité des consonnes impliquées et de la distance entre ces consonnes.

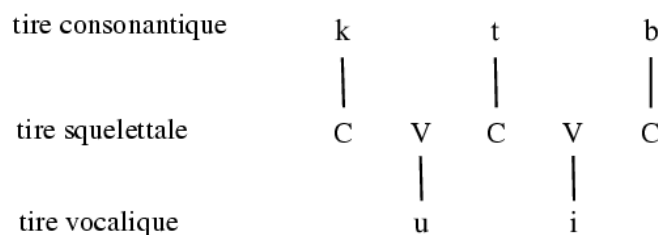


Figure 17 : Représentation du phénomène d'affixation dans la morphologie non-concaténative de l'arabe (extrait de Frisch et al., soumis). La tire consonantique constitue la racine verbale. Les éléments de la tire vocalique sont insérés dans la racine par ancrage à la tire squelettale CV et donnent ainsi lieu à la forme verbale effective.

Les auteurs montrent qu'un modèle assimilant les contraintes phonotactiques de co-occurrence à des contraintes de similarité graduelles rend tout à fait compte de ces données. Partant du principe que la description des phonèmes en termes de traits (acoustiques ou articulatoires) est ancrée non seulement dans la phonologie mais aussi dans la phonétique, cette proposition va donc dans le sens d'un rapprochement de ces deux domaines connexes. Il est par ailleurs tout à fait probable que ce modèle puisse rendre compte des contraintes phonotactiques propres à d'autres langues, par exemple des contraintes de co-occurrence de phonèmes qui seraient sans rapport avec des problèmes de morphologie. Il est en outre possible d'envisager, bien que cette question ne soit pas abordée dans les travaux de Frisch, que cette contrainte de similarité puisse intervenir dans les procédures de syllabation de la chaîne parlée -au même titre que le Principe de sonorité mais à un niveau qui ne serait pas celui de la représentation sous-jacente- tout en reposant sur des caractéristiques effectives des phonèmes mis en jeu dans le processus de syllabation.

<sup>29</sup> Qui ont un ou plusieurs traits en commun.

### 2.2.3. Influences supplémentaires sur la syllabation

On notera qu'un certain nombre de phénomènes conduisent à compliquer la procédure de syllabation qui, finalement et contrairement à ce que l'on peut avoir tendance à penser, ne se résume pas simplement à un processus d'organisation des phonèmes en une structure d'ordre supérieur et qui relèverait uniquement de procédures d'organisation de chaînes phonémiques abstraites. On rappellera en cela la position de Cornulier (1986) qui met en garde contre cette tendance. Parmi ces influences supplémentaires, on compte notamment la structure morphologique des mots et l'accent.

La structure morphologique semble contribuer considérablement au mécanisme de syllabation. Nous avons vu, avec la présentation du *Maximum Onset Principle*, que ce principe contraint un nombre maximal de consonnes intervocaliques à se regrouper à l'attaque de la syllabe suivante. Ainsi, dans une suite CVCCVC composée de trois consonnes médianes, ce principe devrait conduire, si  $C_1C_2$ ,  $C_2C_3$  et  $C_3V$  sont des séquences phonotactiquement légales, à dériver une structure syllabique CV #  $C_1C_2C_3VC$ . Il existe cependant en français des mots qui, bien que proches sur le plan des phonèmes qui les constituent, présenteront certainement des structures syllabiques différentes au niveau de leur prononciation. C'est par exemple le cas des mots 'distribuer' ([distribʁe]) et 'déstresser' ([destrese]). Ces deux mots sont constitués, entre leurs deux premières voyelles, d'une suite identique de 3 consonnes : /str/. Or un locuteur natif proposerait certainement de découper 'distribuer' en [dis # tribʁe] avec une frontière syllabique entre /s/ et /tr/, dissociant par conséquent -et en désaccord avec le MOP<sup>30</sup>- les 3 consonnes du groupe médian par le biais d'un isolement du /s/ en fin de première syllabe. Au contraire, et conformément au MOP, le mot 'déstresser' serait probablement découpé en [de # strese], regroupant ainsi l'ensemble des 3 consonnes à l'attaque de la seconde syllabe. La structure morphologique des mots peut donc contribuer notablement à la structuration syllabique d'une chaîne de phonèmes, du moins pour certains groupes de consonnes.

Dans la section 2.1.1 de ce chapitre, nous avons remarqué que la syllabe constituait une composante à part entière des études sur l'intonation et l'accentuation en donnant un exemple du rôle qu'elle pouvait jouer dans le choix du pattern accentuel d'un mot. Il semble cependant que cette relation entre syllabe et accent puisse se définir également dans le sens inverse. Certaines

<sup>30</sup> Cette discordance entre la syllabation et la structure déterminée par le MOP peut sans doute s'expliquer par les caractéristiques du groupe /st/. Il nous semble en effet raisonnable de poser la question du statut effectif de ce groupe de consonnes en termes de 'légalité' (cf. Section 2.2.2.2 pour une discussion de cette notion).

configurations accentuelles peuvent en effet déterminer la structuration syllabique d'une séquence de segments. Par exemple, la durée d'une voyelle contribue au rattachement de la consonne qui lui succède à la syllabe qui contient cette voyelle ou à la suivante. L'évolution de la hauteur fondamentale de la voix constitue est également liée à la structuration syllabique. Grønnum (1999) met en évidence des configurations de  $F_0$  qui dépendent de la structure syllabique des séquences, illustrant ainsi le lien étroit entre structure syllabique et intonation. Treiman & Danis (1988) montrent par exemple qu'en anglais, les consonnes intervocaliques ne sont ambisyllabiques que si la syllabe qui la suit est faible (non-accentuée). Au contraire, si la seconde syllabe d'une séquence bisyllabique est forte (accentuée), la consonne médiane se situe sans ambiguïté à l'attaque de la seconde syllabe.

### 3. Procédures de syllabation et contraintes phonotactiques

La section précédente nous a permis de passer en revue diverses approches proposées en phonologie pour rendre compte des procédures de structuration syllabique de la chaîne parlée. Notre objectif n'est cependant pas d'étudier la validité de chacune de ces approches mais plutôt d'illustrer en quoi chacune des études comportementales que nous avons citées dans les sections 2.2.3.2 et 2.2.3.3 peut se rattacher préférentiellement à l'une ou l'autre des conceptions théoriques. Nous avons présenté deux approches de la syllabation, chacune constituant à notre avis un point de référence pour les données perceptives illustrant l'interprétation fournie par chacun des auteurs. La première approche envisage la syllabe comme une entité hiérarchiquement structurée qui serait dérivée de règles (ou de contraintes) s'appliquant en partie indépendamment des spécificités acoustiques ou phonétiques des sons impliqués (OOP, MOP, OT). La seconde approche conçoit la syllabe comme une unité qui serait directement déterminée par les caractéristiques propres des phonèmes (principe de sonorité). En dépit des critiques portées à l'encontre des diverses échelles de sonorité proposées et du caractère supposé sous-jacent de ce concept de sonorité, nous avons envisagé que le concept de *similarité* proposé par Frisch (1996) pourrait tout à fait refléter un indice similaire à ceux qui sont envisagés dans les diverses échelles de sonorité (ou de consonanticité) proposés ; indice qui s'ancrerait par contre directement dans une description phonétique et non sous-jacente des segments impliqués.

Un point essentiel et commun à toutes les approches ayant tenté de rendre compte de la syllabation sans faire reposer leurs décisions sur les caractéristiques des phonèmes est qu'en dernière analyse, l'application de ces principes dépend essentiellement de la légalité phonotactique des séquences de segments. En effet, nous avons vu dans la description des

principes d'organisation syllabique que sont l'*Obligatory Onset Principle* et le *Maximum Onset Principle* que leur application dépend en grande partie de ces contraintes phonotactiques. Ces principes s'appliquent à condition que les séquences de phonèmes qu'ils manipulent ne soient pas phonotactiquement illégales. Or cette notion de légalité phonotactique, contrairement au concept de syllabe, peut en partie refléter des contraintes linéaires concernant les caractéristiques propres d'une suite de deux segments telles qu'elles sont définies par le principe de sonorité ou par le concept de similarité. Le point que nous souhaitons développer plus particulièrement dans ce chapitre est donc la question du rapport de ces contraintes phonotactiques à la structuration syllabique. Il est par conséquent essentiel de développer une analyse des liens qui existent entre l'entité hiérarchique 'syllabe' et les contraintes linéaires exprimées dans les régularités phonotactiques de la langue. L'objet de cette section est donc de déterminer (1) ce que l'on entend par légalité phonotactique, (2) de comprendre le lien qui existe entre ce phénomène et la procédure de syllabation et (3) de poser la question de la nécessité d'avoir recours au concept de syllabe pour interpréter les données obtenues dans les tâches perceptives. Nous fournissons donc ici une réflexion sur les relations entre syllabe et légalité phonotactique afin de confirmer cette similarité des processus observés dans les effets mis en évidence par McQueen (1998) et Vroomen & De Gelder (1999).

### 3.1. La notion de légalité phonotactique

Lorsque l'on parle de contraintes phonotactiques, on distingue classiquement deux types de configurations : légales et illégales. Une configuration phonotactiquement illégale consiste en une suite de segments (au minimum deux) qui ne peuvent pas apparaître dans la langue considérée, cette impossibilité se limitant parfois à une position bien précise à l'intérieur des mots ou des syllabes. Par exemple, la séquence /tʎ/ est dite illégale en français. Elle est *illégale* parce qu'elle n'apparaît pas en début de mot. En effet, le groupe consonantique /tʎ/ est attesté en français en position médiane de mot (par exemple dans 'atlantique' ou 'atlas'). Par contre, ce groupe n'apparaît jamais en début de mot ; et tout locuteur natif considérerait la suite /tʎa/ comme *mal formée* en français. On considère en général que cette illégalité provient d'une règle qui interdirait de regrouper deux consonnes coronales à l'intérieur d'une même syllabe. De même, en néerlandais (et probablement aussi en français), /mr/ constitue selon McQueen (1998) une séquence de consonnes phonotactiquement illégale en début de mot. Une séquence de phonèmes illégale serait donc une suite de segments qui ne peut pas apparaître dans certaines positions (ici en début de mot ou de syllabe).

### 3.1.1. Suites de consonnes : groupes et agrégats

Pulgram (1965) insiste sur la nécessité de distinguer, parmi les séquences de consonnes, ce qu'il nomme 'groupe consonantique' et 'agrégat consonantique'. Cette distinction permet d'opposer les séquences de consonnes en fonction de leur appartenance syllabique. Les 'agrégats consonantiques' constitueraient des séquences tautosyllabiques (qui sont prononcées à l'intérieur d'une même syllabe). Au contraire, les 'groupes consonantiques' constitueraient des séquences de consonnes hétérosyllabiques, chaque segment étant alors rattaché à une syllabe différente. Les *agrégats* seraient donc tautosyllabiques alors que les *groupes* seraient hétérosyllabiques. Cette distinction pourrait s'appliquer aussi bien à des séquences de consonnes différentes (par exemple, /tr/ est tautosyllabique alors que /kt/ est hétérosyllabique) qu'à une même séquence prononcée de deux manières différentes (on se rappellera par exemple les différences de syllabation de la suite /st/ dans 'distribuer' et 'déstresser').

### 3.1.2. Groupe légal vs. groupe tautosyllabique

La question qui nous importe ici est de déterminer dans quelle mesure ces deux paramètres (légalité et tautosyllabité vs. illégalité et hétérosyllabité) sont réellement indépendants l'un de l'autre. Il faut évidemment, pour être en mesure de répondre à cette question, se limiter à des situations dans lesquelles ni la structure morphologique ni les phénomènes accentuels (dont on a vu qu'ils contribuent à la structuration d'une chaîne de segments en syllabes) ne modulent ces phénomènes. Nous souhaitons donc poser la question des relations entre légalité phonotactique et syllabation dans un cas bien précis qui ne peut évidemment pas se généraliser à l'ensemble des situations possibles mais qui suffit à notre avis pour mettre en rapport les données obtenues dans les tâches de segmentation lexicale citées. Il est donc important de signaler ici que la réflexion que nous souhaitons engager est en grande partie éloignée de ce qu'est réellement le processus de génération d'une forme de surface tel qu'il est modélisé en phonétique et phonologie. Nous nous situons d'emblée dans une situation très précise : ce phénomène serait conçu comme relevant essentiellement d'un processus d'association d'une chaîne de segments abstraits avec pour seules influences les mécanismes d'organisation hiérarchique en syllabe et de structuration séquentielle à partir des contraintes de sonorité, de similarité ou des régularités phonotactiques.

## 3.2. Qu'est-ce qu'une séquence phonotactiquement légale ?

Si le processus de syllabation repose en grande partie sur le respect des contraintes phonotactiques de la langue, il est nécessaire d'être en mesure de déterminer ce qu'est

effectivement une séquence phonotactiquement légale. Il est cependant difficile de trouver une définition claire de la légalité phonotactique. En effet, cette notion dépend en grande partie du modèle théorique adopté par les auteurs. Nous présentons ici deux positions différentes qui permettent de définir ce qu'est une séquence légale dans la langue.

### 3.2.1. *Toute séquence de phonèmes attestée à l'attaque syllabique*

L'une des définitions possibles de la légalité phonotactique est de considérer comme légale toute séquence de phonèmes qui est possible en début de syllabe dans la langue (Hooper, 1972). Une autre définition possible est de considérer comme légale toute séquence attestée à l'initiale de mot. C'est la définition utilisée par McQueen (1998). Une séquence illégale serait donc constituée de segments qui ne peuvent pas se regrouper à l'initiale de syllabe ou de mot. En réalité, ces deux notions sont intimement liées. C'est sur la mise en évidence de ce lien entre début de mot et début de syllabe que Dell (1995) développe son analyse des groupes de consonnes dans la langue française.

#### 3.2.1.1. Début de syllabe et début de mot

Dell (1995) adopte le raisonnement suivant : toute séquence apparaissant en début de mot constitue également une attaque de syllabe lorsque le mot est prononcé isolément. On peut appliquer le même raisonnement au problème des séquences prononçables en fin de syllabe. La distribution des séquences attestées en début de mot constitue par conséquent un sous-ensemble de celle des séquences attestées en début de syllabe. De même, l'inventaire des séquences pouvant apparaître en position finale de mot constitue un sous-ensemble de celui des séquences apparaissant en fin de syllabe. Dell (1995) choisit donc de recenser les séquences de consonnes qui sont attestées en début et en fin de mot afin de déterminer l'inventaire des groupes de consonnes qui sont en mesure de générer des syllabes 'bien formées'. Selon lui, la distribution des groupes de consonnes constituant des attaques syllabiques bien formées doit permettre de rendre compte des contraintes phonotactiques de la langue (i.e. du caractère légal ou illégal des séquences). Il est évidemment possible que d'autres groupes de consonnes puissent constituer par exemple des attaques de syllabe bien formées mais n'apparaissent pas dans l'inventaire des initiales de mot, cet inventaire constituant seulement une *approximation* des groupes de consonnes qu'il est possible de trouver en attaque ou en coda syllabique. Par conséquent, tout groupe de consonne attesté en début de mot constitue une attaque bien formée alors que l'inverse n'est pas nécessairement vrai : un groupe de consonnes n'apparaissant pas en début de mot ne constitue pas obligatoirement une attaque de syllabe mal formée. L'inventaire des groupes qui ne sont pas attestés en début de mot n'est donc pas assimilable à celui des groupes formant des



attaques de syllabe mal formées (i.e. illégales) ; il reste cependant possible -à notre avis- d'estimer l'inventaire des groupes illégaux en considérant que ceux qui ne sont pas observés en début de mot constituent probablement des attaques mal formées.

#### 3.2.1.2. Présentation des données

A partir de cette relation entre la distribution des groupes de consonnes en début de mot et de syllabe, Dell (1995) effectue un comptage des groupes de consonnes attestés en début de mot dans plusieurs corpus de mots isolés de la langue française (notamment Brulex, Content et al., 1990). Selon lui, tout groupe de consonnes attesté en début de mot constitue une attaque de syllabe bien formée. C'est donc l'existence d'un groupe de consonnes en début de mot qui refléterait sa légalité. A partir de l'analyse qu'il fait des différents *corpora* de son étude, il regroupe l'ensemble des groupes de consonnes recensés en 12 classes dépendant de leurs caractéristiques phonétiques mais aussi de ses intuitions concernant la bonne forme (*well-formedness*) ou légalité des séquences. Cette catégorisation permet déjà de voir que Dell (1995) ne conçoit pas le phénomène de légalité comme une distinction binaire entre d'une part des groupes légaux qui pourraient être prononcés à l'attaque d'une syllabe et d'autre part des groupes illégaux ne pouvant pas se regrouper à l'attaque de syllabe. En effet, parmi les séquences attestées en début de mot dans ses corpora, il propose de classer les divers types de groupes de consonnes en deux classes : la classe correspondant aux groupes donnant lieu à des syllabes bien formées (b, c et d) et celle donnant lieu à des attaques de syllabe bien formées *mais* 'déviants' (e-k). Il existerait donc en réalité 3 classes de groupes consonantiques : les groupes légaux, les groupes illégaux et les groupes légaux mais déviants.

#### 3.2.1.3. La notion de déviance

On peut légitimement se demander ce que signifie cette notion de 'déviante'. Il est clair que la position de Dell (1995) n'est pas de fournir une définition de l'illégalité phonotactique. Il met justement en garde contre cette attitude qui consisterait à conclure, à partir de l'absence d'un groupe dans l'inventaire, à son caractère illégal : il est impossible de savoir si cette absence est un hasard (*accidental gap*) ou est justifiée par les régularités de la langue (*principled gap*). Il reste cependant qu'une position qui consiste à admettre que tout groupe attesté en début de mot constitue une attaque de syllabe bien formée devrait logiquement considérer ces groupes 'déviants' comme bien-formés. Ce qui n'est pas entièrement le cas. En réalité, nous pensons que cette classification dans une catégorie sous le terme de groupes 'déviants' pourrait être liée à la fréquence des groupes de consonnes dans la langue. La notion d'attaque syllabique bien-formée correspondrait alors à quelque chose de plus compliqué qu'un simple recensement des groupes

de consonnes attestés en début de mot et nécessiterait une analyse statistique du lexique pour obtenir des données plus fiables.

### 3.2.2. *Toute séquence de phonèmes fréquente*

Il serait donc possible, si notre proposition est valide, de définir la légalité d'une séquence en termes de fréquence dans la langue, ce qui impliquerait qu'une séquence peut être plus ou moins légale. C'est la définition qu'en donne Altmann (1990) :

Phonotactic constraints reflect the different probabilities of co-occurrence of any two or more phonemes.

Il apparaît que les propositions de Frisch et al. (soumis; Frisch, 1996) concernant le rôle de la similarité pour rendre compte de la légalité des suites phonémiques mettent en évidence un lien étroit entre le concept de similarité et la fréquence des séquences dans la langue. Ce lien refléterait une tendance à juxtaposer plus fréquemment des séquences de phonèmes présentant peu de similarité en termes de traits distinctifs. La légalité phonotactique pourrait donc relever d'un phénomène de similarité des phonèmes entre eux, similarité qui se manifesterait par des différences de fréquence d'utilisation de ces séquences dans la langue. Il serait donc possible de déterminer la légalité phonotactique d'une suite de segments à partir de leur taux de similarité ou de la fréquence d'occurrence de cette suite de phonèmes dans la langue.

### 3.3. Syllabation, contraintes phonotactiques et segmentation lexicale

Les travaux présentés dans le Chapitre 2 (Section 2.2.3) concernant le rôle éventuel des régularités phonologiques dans la segmentation du signal de parole en mots (McQueen, 1998; Vroomen & de Gelder, 1999) fournissent des interprétations alternatives de données comportementales qui sont à notre avis similaires du point de vue des variables manipulées et relèvent de l'opposition entre approche centrée sur la structure hiérarchique de la syllabe et approche centrée sur les caractéristiques des suites de segments. Ces deux interprétations font en effet référence d'une part à un rôle éventuel de la segmentation syllabique (Vroomen & de Gelder, 1999) et, d'autre part, à une segmentation du signal de parole à partir d'indices phonotactiques (McQueen, 1998). Or, comme nous l'avons vu au cours de la description des principes de structuration syllabique (section 2.2), syllabation et contraintes phonotactiques ne sont pas réellement indépendantes l'une de l'autre. L'illégalité phonotactique d'une séquence détermine directement le rattachement des phonèmes qui la constituent à la hiérarchie syllabique. La structure syllabique est donc -en partie- une conséquence des régularités phonotactiques. Les

interprétations alternatives proposées par les auteurs de ces études (Vroomen & de Gelder, 1999; McQueen, 1998) sont en fait déterminées par le cadre théorique qui est choisi comme référence. Vroomen & De Gelder (1999) se fondent sur le concept hautement hiérarchique de syllabe pour analyser leurs données alors que McQueen (1998) fait appel à la notion de contraintes phonotactiques pour interpréter les résultats qu'il obtient. On notera que dans une dernière expérience, McQueen (1998) tente de dissocier -sans succès cependant- les rôles respectifs des contraintes phonotactiques et de la syllabation en demandant au locuteur qui enregistre les stimuli de prononcer les séquences médianes, qu'elles soient légales ou pas, de deux manières : soit regroupées à l'attaque syllabique, soit avec une frontière syllabique insérée entre les deux phonèmes. Si aucun effet de la syllabation n'émerge indépendamment de celui de la légalité phonotactique dans la tâche de segmentation lexicale, c'est probablement en raison du lien étroit qui existe entre ces deux phénomènes. Il est probable que la tâche de production d'un groupe illégal à l'intérieur d'une syllabe unique soit particulièrement difficile à accomplir. De même, il est peut-être délicat de contrôler consciemment la prononciation d'un groupe légal dans deux syllabes différentes. Cette difficulté éventuelle de contrôle des phénomènes articulatoires par le locuteur pourrait expliquer l'absence d'effets indépendants et serait déterminée par le lien étroit qui unit contraintes phonotactiques et syllabation.

Nous affirmons donc que les données comportementales obtenues par Vroomen & De Gelder (1999) et par McQueen (1998) sont analysables en termes de mécanismes cognitifs similaires qui reposeraient soit sur une représentation syllabique de la chaîne de phonèmes, soit sur la détection de frontières phonotactiques. Du fait de l'intrication entre régularités phonotactiques et syllabation, il semble difficile de dissocier ces deux composantes du traitement cognitif mis en œuvre par les auditeurs. On pourra objecter que l'expérience réalisée par Vroomen & De Gelder (1999) n'a pas recours à des séquences considérées comme phonotactiquement illégales et que, par conséquent, ces résultats ont nécessairement un lien avec le processus de syllabation. En effet, les auteurs comparent des séquences Consonne-Voyelle dans lesquelles la consonne est nécessairement prononcée à l'attaque syllabique et des séquences Consonne-Consonne qui ne sont pas considérées comme illégales en néerlandais mais induisent néanmoins une syllabation de la première consonne en coda syllabique. Cette proposition n'est cependant valide que si l'on peut affirmer que la légalité est un phénomène binaire, auquel cas l'on est nécessairement amené à mettre en œuvre des procédures spécifiques de syllabation (par exemple le *Maximum Onset Principle*) afin de déterminer la structure syllabique d'un énoncé composé de séquences phonotactiquement légales. En réalité, si l'on se réfère aux travaux de Frisch et al. (soumis), il semble que la légalité d'une chaîne de segments constitue plutôt un

continuum sur lequel chaque séquence pourrait se situer. Il serait alors beaucoup plus difficile de dissocier légalité et syllabation, et la distinction de structure syllabique qui est présentée par Vroomen & De Gelder (1999) pourrait être assimilée à une position différente des séquences Consonne-Consonne et Consonne-Voyelle sur un continuum de légalité, les groupes occlusive-fricative étant moins légaux que les suites occlusive-voyelle.

En raison de cette indissociabilité des effets liés à une représentation syllabique de la chaîne de phonèmes ou à la localisation de frontières phonotactiques, nous considérerons désormais que les effets observés dans les expériences présentées (McQueen, 1998; Vroomen & de Gelder, 1999) reposent probablement sur un mécanisme commun qui correspondrait à une segmentation du signal de parole reposant sur des contraintes de nature phonologique (que ces contraintes reposent sur les principes de syllabation ou sur les régularités phonotactiques). Nous choisissons par conséquent d'appréhender la question du rôle des contraintes phonologiques dans les processus cognitifs de segmentation lexicale en nous affranchissant de chercher à distinguer entre structure syllabique et légalité phonotactique. Pour cela, nous considérerons par la suite les diverses conditions expérimentales mises en œuvre dans les expériences comportementales sous l'angle des caractéristiques linéaires des consonnes impliquées plus qu'en fonction de leur tautosyllabité avérée. Nous adopterons le même principe dans le chapitre suivant lorsque nous présenterons une analyse distributionnelle des groupes de consonnes dans le lexique. Evidemment, nous garderons en permanence à l'esprit que ces caractéristiques phonétiques déterminent en grande partie leur rattachement à la hiérarchie syllabique mais aussi que des groupes présentant des caractéristiques similaires peuvent être syllabés différemment. Cette absence de référence explicite à des structures hiérarchiques nous permettra cependant de proposer des interprétations différentes des effets observés en nous donnant l'occasion de considérer les variables impliquées sous un angle purement linéaire. Les résultats obtenus par McQueen (1998) et par Vroomen & De Gelder (1999) pourraient notamment trouver leur source dans des processus de calcul statistique similaires à ceux qui sont proposés par Saffran, Newport, & Aslin (1996; cf. chapitre suivant pour une justification plus approfondie) plutôt que dans l'intervention de réelles connaissances sur les régularités phonologiques de la langue. Des travaux présentant des simulations connexionnistes du phénomène de syllabation d'une chaîne de phonèmes mettent par ailleurs en évidence la possibilité qu'ont ces modèles de faire émerger des structures syllabiques à partir d'une chaîne linéaire de phonèmes. Ces résultats mettent en évidence la capacité des modèles connexionnistes à développer des propriétés émergentes conformes aux structures syllabiques observées dans les langues aussi bien dans le cadre de la phonologie (Laks, 1995) que de la psycholinguistique développementale (Vroomen, van den

Bosch, & de Gelder, 1998). Typiquement, un modèle connexionniste -ne disposant pas d'informations spécifiques et abstraites concernant ce qu'est une structure syllabique- ne peut que se fonder sur des indices statistiques pour développer des propriétés émergentes ne relevant pas explicitement de principes statistiques (comme la structure syllabique). Il semble donc clair, du fait de cette capacité des modèles connexionnistes à 'apprendre' ce qu'est une syllabe, que des informations probabilistes peuvent être disponibles pour la segmentation et la structuration syllabiques. Il apparaît, à la lumière du lien étroit qui peut exister entre mécanismes de syllabation, contraintes phonotactiques et fréquence d'occurrence des phonèmes impliqués que les interprétations proposées respectivement par McQueen (1998) et Vroomen & De Gelder (1999) pourraient reposer sur des processus cognitifs qui n'auraient pas directement recours à des connaissances phonologiques mais seraient plutôt liés à la fréquence d'occurrence des paires de segments dans la langue. Il est donc nécessaire de conduire une analyse de la fréquence des groupes de consonnes dans le lexique afin de poser les éventuels fondements d'une critique des interprétations proposées par ces auteurs en vérifiant la légitimité du lien qui est envisagé entre régularités phonologiques et fréquence d'occurrence dans la langue.

## Résumé

Dans la première partie de ce chapitre, nous avons présenté une introduction aux problématiques rencontrées dans le domaine de la phonologie. Cette introduction nous a permis de mettre en évidence le rôle de représentations hiérarchiques dans les modèles proposés et notamment de celui de 'syllabe'. Nous avons alors entamé une réflexion sur ce qu'est la syllabe et avons insisté sur la distinction qui peut être faite entre sa structure en constituants et certains paramètres linéaires qui pourraient permettre de découper une séquence de phonèmes en syllabes sans avoir nécessairement recours à une conception hautement hiérarchique. L'un des points essentiels qu'il nous a semblé important de soulever est le lien intime qui peut exister entre syllabe et contraintes phonotactiques. Ce concept de contraintes phonotactiques est cependant conçu selon des angles divers par le biais desquels il est possible de concevoir l'opposition entre suite légale et illégale comme une distinction binaire (est légale une séquence attestée dans la langue) ou stochastique (auquel cas il serait possible de définir la légalité comme une échelle continue allant du moins légal au plus légal et qui se manifesterait par la fréquence d'occurrence des séquences dans la langue). Si procédures de syllabation, contraintes phonotactiques et fréquence d'occurrence des segments

sont intimement liées, les effets avancés comme des preuves du recours à des connaissances sur les régularités phonologiques de la langue dans les processus de segmentation lexicale pourraient s'expliquer différemment. Ceci nous conduit à affirmer la nécessité de conduire une analyse distributionnelle des paires de segments dans la langue afin de poser les fondements d'une analyse pertinente des données comportementales présentées.