



Université de Bretagne-Sud

UFR SSI – Master DTIC



Traitement Automatique des Langues

TRAVAUX DIRIGES

Enseignant

Jean-Yves ANTOINE

(Université François Rabelais)

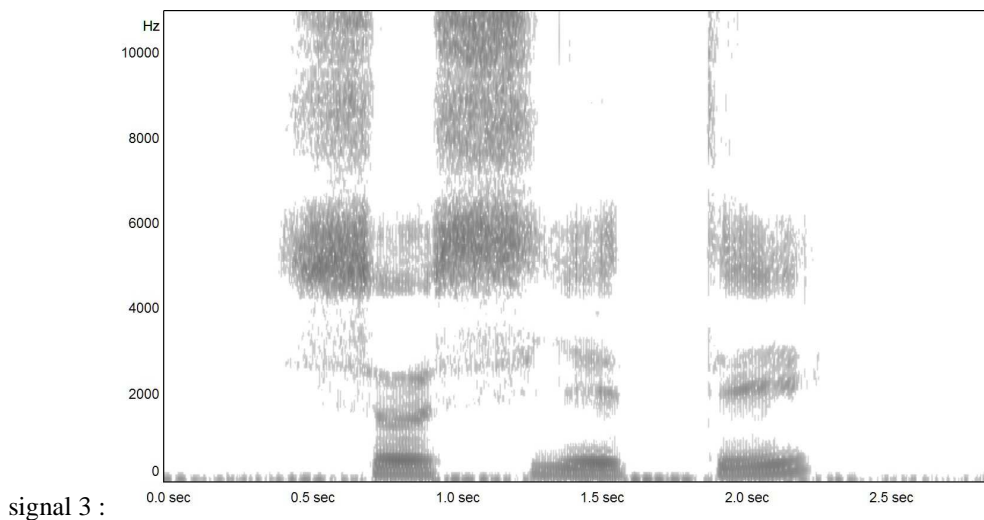
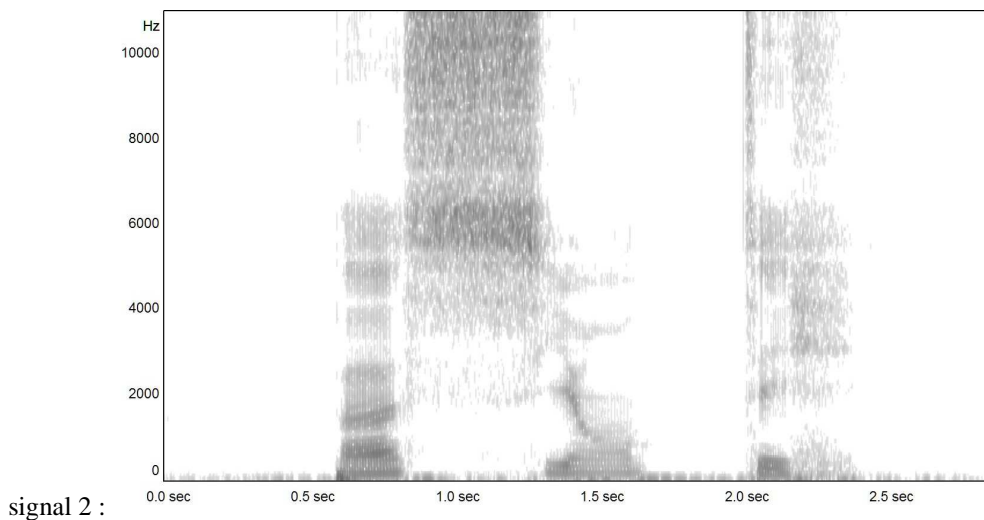
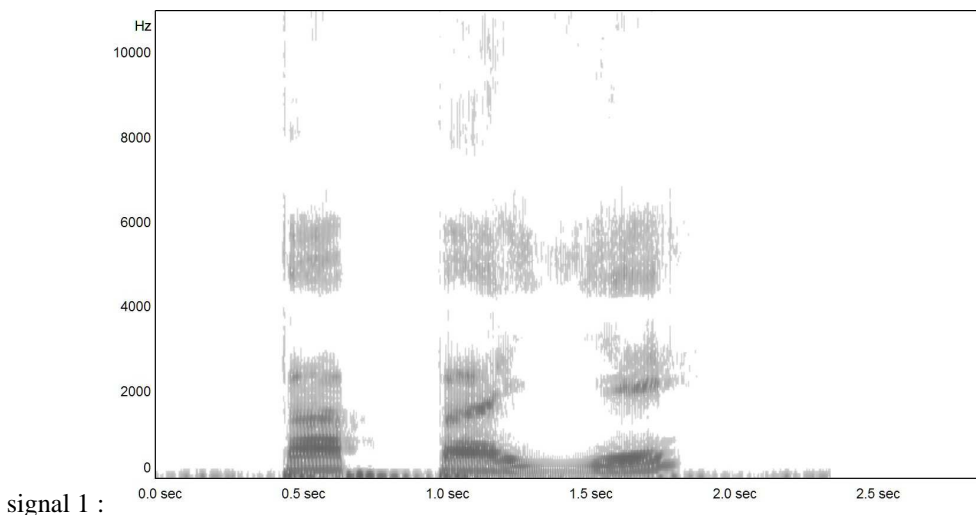
Phonétique

Exercice 1 : Transcription phonétique

- 1) Donnez la transcription dans l'alphabet phonétique international des mots suivants : *suave, tien, ouate, souhait, tueur, gageure, plier, joie.*
- 2) Sauriez-vous retrouver les mots anglais dont voici la transcription phonétique : /spitʃ/, /dʒestʃə/, /lesn/

Exercice 2 : Lecture de spectrogramme

Rattacher les spectrogramme ci-dessous aux mots suivants : a – *société* b – *patienter* c – *papayer*

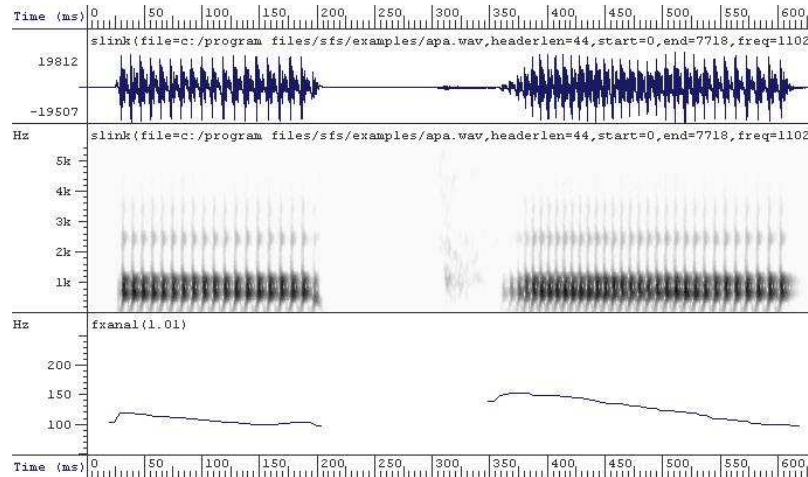


Exercice 3 : Analyse du signal de parole et lecture de spectrogramme

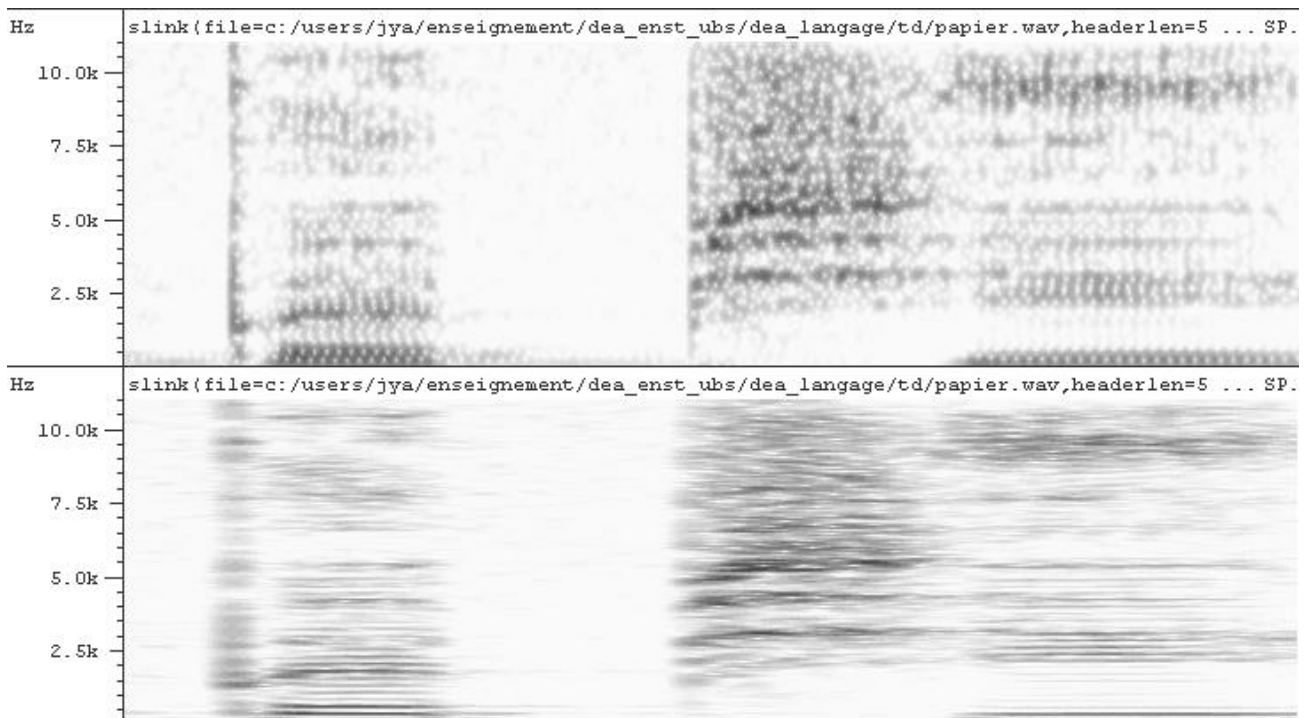
Dans cet exercice, nous allons nous intéresser aux caractéristiques du signal de parole tel qu'il peut être caractérisé par l'analyse de son énergie, son spectre ainsi que sa fréquence fondamentale (pitch) en cas de voisement.

1 — On considère le signal de parole ci-dessous, dont on donne précisément (de haut en bas) l'évolution au cours du temps de l'énergie, du spectre (spectrogramme) et enfin de la fréquence fondamentale. Ce signal de parole est constitué de trois phonèmes successifs. Le premier et le dernier phonèmes correspondent au /a/.

Pouvez-vous donner la liste des phonèmes qui sont susceptibles de correspondre au phonème du milieu ? On justifiera la réponse aussi précisément que possible.



2 — On considère maintenant le signal de parole ci-dessous, dont on donne deux types de spectrogramme. Déterminez lequel des deux spectrogrammes correspond à une analyse à bande étroite. Justifiez votre réponse.

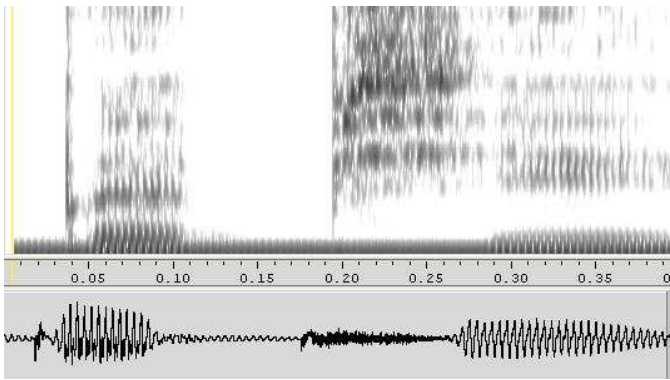


3 — On considère maintenant quatre signaux de parole, pour lesquels on donne à chaque fois (de haut en bas) le spectrogramme et l'enveloppe de l'énergie (cette dernière information est légèrement décalée temporellement par rapport au spectrogramme). Ces signaux correspondent, par ordre alphabétique, aux mots suivants :

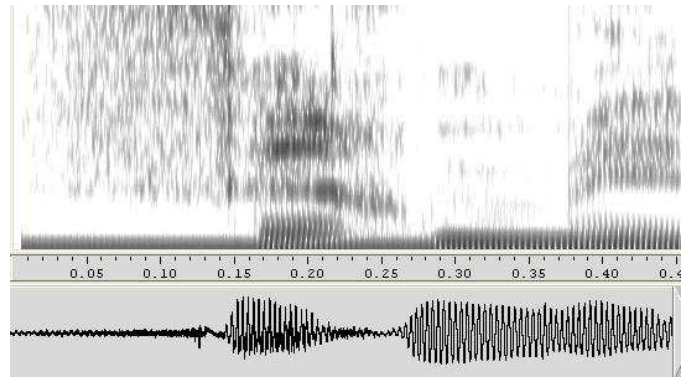
- | | |
|------------|------------|
| 1) changer | 2) fermer |
| 3) papier | 4) reprend |

Associez chacun des signaux ci-dessous au mot qui lui correspond. Expliquez à chaque fois quels sont les indices qui orientent votre décision.

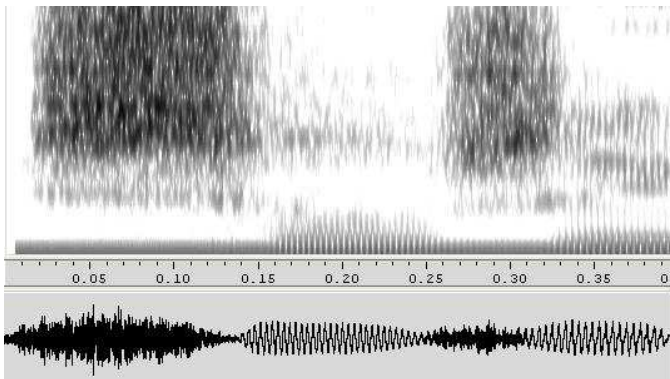
Signal A



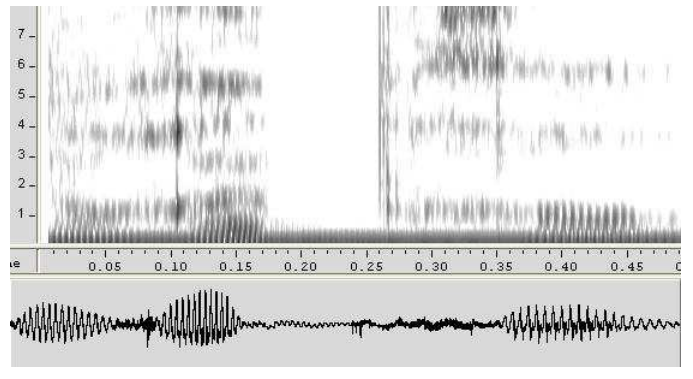
Signal B



Signal C



Signal D

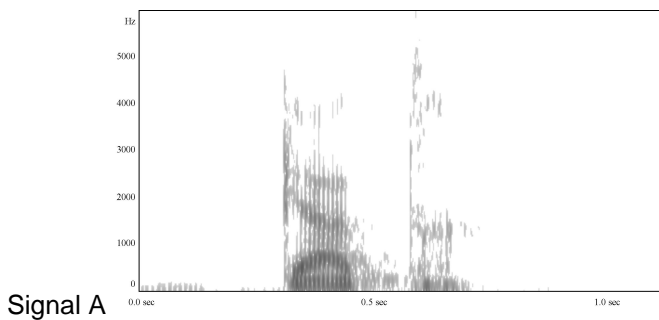


Exercice 4 : Lecture de spectrogramme

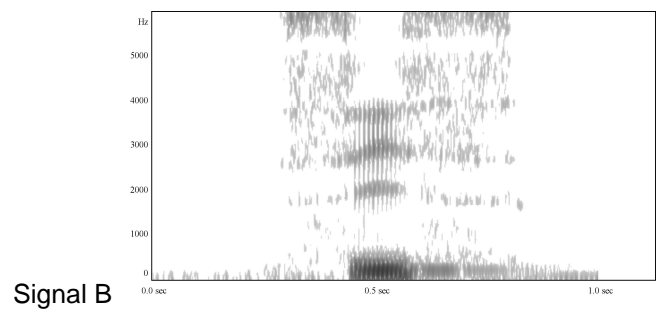
Certains serveurs vocaux limitent le dialogue entre l'homme et la machine à une navigation par sous-menus où le locuteur doit prononcer à chaque étape un numéro correspondant à un choix dans une liste. Par exemple :

- Désirez-vous :
- 1 *quitter l'application*
 - 2 *connaître le solde de votre compte*
 - 3 *connaître vos dernières opérations*
 - 4 *être mis en relation avec une opératrice*

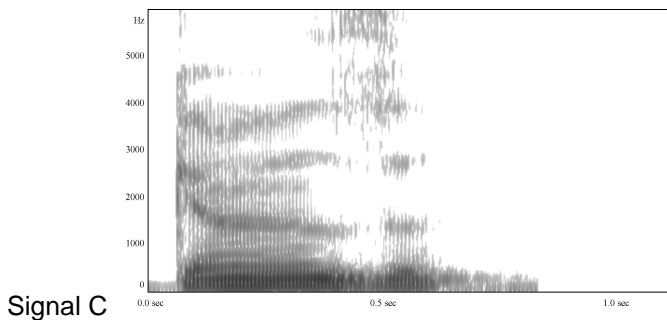
Dans cette interaction assez fruste, la reconnaissance de la parole se limite à l'identification du nombre prononcé par le locuteur. Tâche facile ? C'est ce que nous allons voir dans cet exercice de lecture de spectrogramme, où vous allez jouer le rôle du système de reconnaissance.



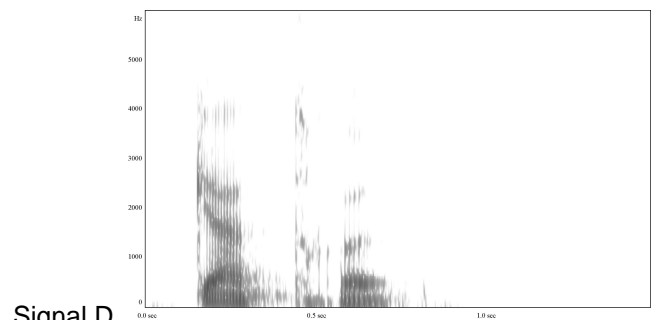
Signal A



Signal B



Signal C



Signal D

Les spectrogrammes de la page précédente correspondent à la prononciation d'un des nombres suivants :

« 0 »

« 3 »

« 4 »

« 6 »

« 15 »

Chaque nombre peut correspondre à 0, 1 ou 2 signaux de parole. Saurez-vous retrouver le nombre correct dans tous les cas ?

Morphologie et syntaxe

Exercice 1 : Morphologie dérivationnelle

En français, il est souvent possible de nominaliser les verbes, c'est à dire de créer un substantif pour dénommer une action. Un des morphèmes les plus productifs pour réaliser cette opération est le suffixe *-ation*. Par exemple : *acceptation, multiplication, généralisation, réparation*.

- 1) Comment analyseriez-vous la dérivation des termes : *soustraction, destruction, instruction*.
- 2) Trouvez au moins deux autres morphèmes permettant de réaliser cette opération de nominalisation en français et donnez des exemples de lexèmes ainsi produits.
- 3) Le morphème *-ure* réalise-t-il cette nominalisation : *rayure, glaçure, gelure, procédure...*
- 4) Les morphèmes que vous avez proposés à la question 2) sont-ils ambigus ou non ?

Exercice 2 : Morphologie compositionnelle et dérivationnelle

Comment analyseriez-vous la construction morphologique (composition, dérivation) des termes suivants.

- *Désintoxication*
- *Indécrottable*
- *Porte-feuille*
- *Cuirassé*
- *Greffe et châtain*

Exercice 3 : Segmentation d'un énoncé en syntagmes

Dans cet exercice, on se place dans une perspective purement syntagmatique. On demande d'identifier les syntagmes composant chacun des énoncés ci-dessous. On ne donnera dans un premier temps que les syntagmes principaux de l'énoncé, c'est-à-dire qu'on ne cherchera pas à identifier, par exemple, le groupe prépositionnel complément du nom d'un syntagme nominal tel que « *la voiture de mon père* ». On donnera ensuite une segmentation plus fine en détail les groupes syntagmatiques ou proposition enchâssés dans les syntagmes.

- (a) *je regarde la montagne enneigée depuis ma fenêtre*
- (b) *Jean peut lui prêter de l'argent très rapidement*
- (c) *l'avion et la voiture sont des modes de déplacement très polluants*
- (d) *toute parole n'est pas bonne à dire*
- (e) *sa mère lui a demandé de téléphoner dès son arrivée*
- (f) *je pense que la personne qui a fait ce forfait est partie dès que l'alarme a sonné*

Exercice 4 : Arbres syntaxiques

Donnez l'arbre syntaxique des énoncés ci-dessous, suivant une approche purement syntagmatique :

- (a) *dépose la clef dans la boîte aux lettres*
- (b) *tous les autres enfants sont partis avec leur mère*
- (c) *Jacques a parlé de ses problèmes avec le juge*
- (d) *le gardien et le directeur confirment que le portail était fermé*
- (e) *la fille que tu regardes est ma sœur*
- (f) *l'homme dont tu m'as parlé a retéléphoné dès ton départ.*

Exercice 5 : Structures syntaxiques

Donner la structure syntaxique, suivant une approche purement syntagmatique à la Chomsky, des énoncés ci-dessous :

- (a) *Tout corps qui est plongé dans un liquide subit la pression d'Archimède*
- (b) *Tous les enfants coururent vers la sortie dès que la cloche libératrice sonna*
- (c) *Le voleur refusa d'indiquer aux enquêteurs l'endroit où était caché le butin*
- (d) *Il règne toujours la même ambiance festive dans les villes espagnoles*
- (e) *Les assiégés eurent à choisir entre une mort honorable et l'infamie*
- (f) *Ils déclarèrent alors préférer mourir et rester dignes*

Exercice 6 : Structures syntaxiques

1 — Donner la structure syntaxique, suivant une approche purement syntagmatique à la Chomsky, des énoncés ci-dessous :

- (a) *Tous les députés sortirent de l'hémicycle dès que le leader nationaliste pris la parole*
- (b) *Les enquêteurs essayèrent de déterminer précisément l'heure à laquelle le suspect quitta son hôtel*
- (c) *Tout individu qui ne respectait pas le couvre-feu était arrêté par la milice*
- (d) *Le même froid dramatique règne sur l'ensemble des régions*
- (e) *La campagne se déroule désormais entre le président actuel et son challenger démocrate*
- (f) *Ils décidèrent finalement de partir sans avoir mangé*

2 — Pouvez-vous donner les structures de dépendances qui correspondraient aux énoncés (a) et (e) ci-dessus ?

3 — Quel type de problème pose à une analyse purement syntagmatique l'énoncé suivant :

- (d') *Sur l'ensemble des régions règne le même froid dramatique*

4 — En quoi une analyse de type LFG résout-elle ce problème ?

Exercice 7 : Rection verbale

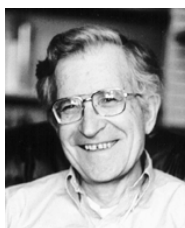
1) Quels sont, parmi les verbes suivants, ceux qui sont transitifs et ceux qui sont intransitifs (ou les deux) :

- (a) *assassiner*
- (b) *coller*
- (c) *patienter*
- (d) *envoyer*

2) Donnez l'ensemble des compléments appartenant à la valence des verbes transitifs, en distinguant les arguments obligatoires des arguments optionnels.

Grammaires formelles

Exercice 1 : Hiérarchie de Chomsky



Que ce soit en informatique (compilation) ou en traitement des langues naturelles, seules les grammaires régulières et hors-contexte définies par Chomsky sont en pratique utilisées. Il n'est cependant pas inutile de connaître l'ensemble de la hiérarchie de grammaires formelles définies par Chomsky. Cet exercice vise précisément à vous la remémorer.

On considère différentes grammaires définies sur le vocabulaire terminal $V_t = \{ a, b \}$ et utilisant comme axiome le symbole non terminal $\langle S \rangle$:

- (A) $\langle S \rangle \rightarrow \langle R \rangle$
 $\langle S \rangle \rightarrow a \langle R \rangle b$
 $\langle R \rangle \rightarrow a$
 $\langle R \rangle \rightarrow b$
- (B) $\langle S \rangle \rightarrow b \langle R \rangle$
 $\langle R \rangle \rightarrow a \langle R \rangle$
 $\langle R \rangle \rightarrow b$
- (C) $\langle S \rangle \rightarrow \langle R \rangle$
 $\langle S \rangle \rightarrow a \langle R \rangle$
 $a \langle R \rangle \rightarrow \langle S \rangle$
 $\langle R \rangle \rightarrow \langle R \rangle b$
 $\langle R \rangle \rightarrow a$
 $\langle R \rangle \rightarrow b$

1. Donnez le type, dans la hiérarchie de Chomsky, de ces grammaires.
2. Quel est le langage généré (forme générale des productions $L\langle S \rangle$) par chacune de ces grammaires ?

Exercice 2 : Chips ou cacahouètes ?

Les grammaires peuvent être utilisées en reconnaissance des formes pour, par exemple, la description de caractères orthographiques. De même, on peut les utiliser pour générer des motifs graphiques. C'est le cas de cette petite grammaire relativement artificielle, définie sur le vocabulaire suivant :

$$V = \{ \epsilon, \exists, \equiv \}$$

La base de règle est la suivante :

- 1) $\langle S \rangle \rightarrow \langle D \rangle \langle MS \rangle \langle F \rangle$
- 2) $\langle MS \rangle \rightarrow \langle M \rangle$
- 3) $\langle MS \rangle \rightarrow \langle M \rangle \langle RS \rangle \langle M \rangle$
- 4) $\langle RS \rangle \rightarrow \langle F \rangle \langle S \rangle \langle D \rangle$
- 5) $\langle F \rangle \rightarrow \exists$
- 6) $\langle D \rangle \rightarrow \epsilon$
- 7) $\langle M \rangle \rightarrow \equiv$

- 1 — Donner trois expressions différentes engendrées par la grammaire, accompagnées de leur arbre de dérivation.
- 2 — Quelle la forme générale des expressions engendrées par la grammaire ?
- 3 — **Plus difficile (inférence de grammaire)** — A quel type de grammaire (suivant la typologie de Chomsky) appartient cette grammaire ? Peut-on engendrer le même langage par une grammaire d'un type plus élevé ? Si oui, donner les règles de cette grammaire.

Exercice 3 : « Bonjour gros grand gras grain d'orge » ou une grammaire récursive

En français, il est possible de juxtaposer plusieurs adjectifs antéposés sans aucun coordonnants. On retrouve ce type de construction dans une petite comptine enfantine (et ne me dites pas que cet exemple est tiré par les cheveux !):

*Bonjour gros grand gras grain d'orge
Quand te dégros-grand-gras-grain-d'orgeras-tu ?
Je me dégros-grand-gras-grain-d'orgerai
Quand tout les gros grands gras grains d'orges
Se dégros-grand-gars-grain-d'orgeront*

On considère donc la pseudo-grammaire ci-dessous, d'axiome $\langle \text{SN} \rangle$ et définie sur le vocabulaire *gros, grand, gras, grain, le* :

- | | | | |
|-----|-------------------------------|---------------|--|
| (1) | $\langle \text{SN} \rangle$ | \rightarrow | $\langle \text{DET} \rangle \langle \text{GN} \rangle$ |
| (2) | $\langle \text{GN} \rangle$ | \rightarrow | $\langle \text{GADJ} \rangle \langle \text{N} \rangle$ |
| (3) | $\langle \text{GN} \rangle$ | \rightarrow | $\langle \text{NC} \rangle$ |
| (4) | $\langle \text{GADJ} \rangle$ | \rightarrow | $\langle \text{GADJ} \rangle \langle \text{ADJ} \rangle$ |
| (5) | $\langle \text{GADJ} \rangle$ | \rightarrow | $\langle \text{ADJ} \rangle$ |

Dans toutes les questions ci-dessous, il est demandé de donner l'arbre de dérivation correspondant à la stratégie de recherche considérée. Dans le cas d'une dérivation se concluant positivement, on donnera en outre l'arbre d'analyse obtenu.

- 1 — Un parseur ne doit pas se contenter d'analyser correctement les énoncés du langage, il doit aussi rejeter les énoncés incorrects. Donnez l'arbre correspondant à un stratégie de recherche descendante en profondeur d'abord sur l'énoncé *un gros gros*. Conclusion ?
- 2 — Donnez l'arbre correspondant à un stratégie de recherche descendante en largeur d'abord sur l'énoncé *un gros gros*. Conclusion ?
- 3 — Donnez l'arbre correspondant à un stratégie de recherche descendante en largeur d'abord sur l'énoncé *un gros grain*. Conclusion ?
- 4 — Donnez l'arbre correspondant à un stratégie de recherche descendante en profondeur d'abord sur l'énoncé *un gros grain*. Conclusion ?
- 5 — Pourriez-vous donner une grammaire qui fonctionne sur les cas ayant posé problème ?

Exercice 4 : Grammaire de parenthésage *adapté de Aho, Sethi & Ullman, Compilateurs, InterEditions*

On considère la grammaire suivante, construite sur le vocabulaire terminal $V_t = \{ a ; (;) ; , \}$ et d'axiome $\langle S \rangle$.

- | | | | |
|-----|---------------------|---------------|---|
| (1) | $\langle S \rangle$ | \rightarrow | $(\langle L \rangle) a$ |
| (2) | $\langle L \rangle$ | \rightarrow | $\langle L \rangle , \langle S \rangle \langle S \rangle$ |

- 1 — Donner les arbres d'analyse des 2 phrases suivantes :
 - a) $(a,(a,a))$
 - b) $(a,((a,a),(a,a)))$
- 2 — Donner la succession des dérivations correspondant à l'analyse de l'énoncé a) par une stratégie ascendante en profondeur d'abord.
- 3 — Donner la succession des dérivations correspondant à l'analyse de l'énoncé a) par une stratégie d'analyse descendante avec dérivation gauche et stratégie de recherche en profondeur d'abord. Que constatez-vous ? Sauriez-vous résoudre ce problème en modifiant la grammaire ?

Exercice 5 : Analyse du français et stratégie de recherche

On considère la grammaire (très limitée) du français d'axiome $\langle S \rangle$ définie ci-dessous :

- | | | | |
|-----|-----------------------------|---------------|---|
| (1) | $\langle S \rangle$ | \rightarrow | $\langle \text{SN} \rangle \langle \text{SV} \rangle$ |
| (2) | $\langle \text{SN} \rangle$ | \rightarrow | $\langle \text{DET} \rangle \langle \text{NC} \rangle$ |
| (3) | $\langle \text{SN} \rangle$ | \rightarrow | $\langle \text{DET} \rangle \langle \text{ART} \rangle \langle \text{NC} \rangle$ |
| (4) | $\langle \text{SV} \rangle$ | \rightarrow | $\langle \text{V} \rangle$ |
| (5) | $\langle \text{SV} \rangle$ | \rightarrow | $\langle \text{V} \rangle \langle \text{SN} \rangle$ |

- 1 — Donner une dérivation gauche possible de l'énoncé *le chien regarde le chat* par cette grammaire. Cette grammaire est-elle déterministe ?
- 2 — Donner un arbre d'analyse possible de l'énoncé *le chien regarde le chat*. Cette grammaire est-elle ambiguë ?
- 3 — Donner la succession des états correspondant à l'analyse de l'énoncé *le chien regarde le chat* suivant les stratégies définies ci-dessous :
 - a) analyse descendante avec stratégie de recherche en profondeur d'abord
 - b) analyse descendante avec stratégie de recherche en largeur d'abord
 - c) analyse ascendante tabulaire (char parser). Donnez de manière graphique l'agenda (ensemble des constituants du chart) final obtenu.
 - d) analyse descendante tabulaire (**algorithme de Earley : hors programme cette année...**)

Exercice 6 : Chart parsing

On considère la grammaire hors-contexte d'axiome <S> définie ci-dessous :

- (1) <S> → <SN> <SV>
- (2) <SV> → <V>
- (3) <SV> → <AUX> <PP>
- (4) <SN> → <DET> <NC>

- 1 — Définissez le lexique correspondant à cette grammaire pour le vocabulaire suivant : {*la, souris, fuit, a*}
- 2 — Donner la succession des états correspondant à l'analyse de l'énoncé *la souris a fuit* suivant une analyse ascendante tabulaire (algorithme coin-gauche).
- 3 — Donner la succession des états correspondant à l'analyse de l'énoncé *la souris a fuit* suivant une analyse descendante tabulaire (algorithme de Earley : hors programme cette année...).

Exercice 7 : Une grammaire ambiguë

On considère la grammaire hors-contexte suivante :

- (1) <S> → <SN> <SV>
- (2) <SN> → <DET> <NC>
- (3) <SN> → <DET> <ADJ> <NC>
- (4) <SV> → <GV>
- (5) <SV> → <GV> <SN>
- (6) <GV> → <V>
- (7) <GV> → <CLITIQUE> <V>

LEXIQUE :

belle	: ADJ
belle	: NC
ferme	: NC
ferme	: V
la	: DET
la	: CLITIQUE
le	: DET
le	: CLITIQUE
masque	: NC
masque	: V

- 1 — Donnez l'arbre d'analyse de l'énoncé « *la belle ferme le masque* ». Cette grammaire est-elle ambiguë ?
- 2 — Etudiez l'influence de cette ambiguïté en donnant la suite des dérivations effectuées lors de l'analyse de l'énoncé :
 - a) suivant une stratégie de recherche ascendante en profondeur d'abord
 - b) suivant une stratégie de recherche descendante en profondeur d'abord

Exercice 8 : Du bon usage du groupe nominal en français

La mise en œuvre d'une grammaire du français relativement couvrante est une tâche de longue haleine qu'on ne saurait réaliser en 2h de TD... Aussi, dans cet exercice, allons-nous intéresser à un phénomène linguistique très précis : le groupe nominal en français. Dans cet exercice, le groupe nominal français peut comporter soit un nom propre, soit un déterminants suivi d'un nom, éventuellement encadré d'adjectifs antéposé ou postposé. Adjectifs et déterminants peuvent en outre faire partie d'un groupe déterminatif ou d'un groupe adjectival. Par exemple :

<i>la bonne soupe</i>	<i>une soupe froide</i>
<i>toute la soupe</i>	<i>une autre soupe</i>

En outre, le groupe nominal peut également comprendre un groupe prépositionnel complément du nom :

la bonne soupe de la grand-mère paternelle de Jean

- 1 — Ecrire une grammaire hors-contexte qui modélise ces groupes nominaux minimaux.
- 2 — Donnez l'arbre d'analyse correspondant au groupe nominal *la très bonne soupe de Jean*
- 3 — Votre grammaire sur-génère-t-elle ? Les dérivations incriminées sont-elles incorrectes d'un point de vue syntaxique ou sémantique ?
- 4 — Combien de règles compterait votre grammaire si vous décidiez de gérer les phénomènes d'accord.
- 5 — On suppose que l'on dispose d'un ensemble de règles permettant de décrire les relatives, de catégorie <SUBREL>. Comment rattacher ces additions à votre grammaire ?

Grammaires d'unification

Exercice 1 : Du bon usage du groupe nominal en français (suite)

On considère la grammaire du groupe nominal construite dans le TD1. Pouvez-vous augmenter cette grammaire à l'aide de structures de traits pour gérer les phénomènes d'accord ? Donnez ensuite une traduction de cette grammaire en LFG.

Exercice 2 : Augmentation d'une grammaire hors-contexte

On considère la grammaire hors - contexte suivante :

(1)	<S>	→	<SN> <SV>
(2)	<SN>	→	<DET> <NC>
(3)	<SV>	→	<V>
(4)	<SV>	→	<V> <SN>
(5)	<SV>	→	<V> <SP>
(6)	<SV>	→	<COPULE> <ADJ>
(7)	<SP>	→	<PREP> <SN>

LEXIQUE :

mère	: NC
porte	: NC
porte	: V
enfant	: NC
la	: DET
le	: DET
les	: DET
son	: DET
sur	: PREP
dos	: NC
parle	: V
donne	: V
sortie	: NC
sortie	: ADJ
est	: COPULE
à	: PREP

1. Donnez la succession des états correspondant à l'analyse de l'énoncé *la mère porte son enfant* suivant une analyse descendante avec stratégie de recherche en profondeur d'abord.
2. Donnez l'arbre de dérivation correspondant à cette analyse.
3. Donnez la succession des états correspondant à l'analyse de l'énoncé *la porte donne sur la sortie* suivant une analyse ascendante tabulaire.
4. Donnez l'agenda du chart obtenu à la fin de cette analyse, ainsi que l'arbre d'analyse correspondant à l'énoncé ainsi traité.
5. Donnez quatre exemples d'énoncés appartenant au langage généré par la grammaire qui ne respectent pas la syntaxe du français. On cherchera à trouver des erreurs de nature différente à chaque fois, en expliquant la cause de l'erreur obtenue.
6. Afin de limiter ce genre de surgénérations, il est possible d'augmenter cette grammaire à l'aide de structures de traits (sous-catégorisation, traits morphologiques) permettant de contraindre le déclenchement des règles. Dans cet exercice, cette grammaire attribuée sera une grammaire lexicale-fonctionnelle (LFG).
7. Donnez la liste des traits que vous allez utiliser dans votre grammaire LFG pour contrôler les surgénérations constatées à la question précédente. Donnez également la liste des fonctions syntaxique que vous allez utiliser pour construire la f-structure de vos énoncés.
8. Adaptez la grammaire précédente dans le formalisme LFG afin d'éviter tout surgénération incorrecte sur le vocabulaire considéré dans cet exercice.

Exercice 3 – Grammaire attribuée : traduction en LFG et en TAG

On considère la grammaire attribuée suivante :

- (1) $\langle S \rangle \text{ MOD}=\textit{ind}$ \rightarrow $\langle \text{SN} \rangle \text{ NB}=\textit{X PERS}=\textit{Y GR}=\textit{Z}$ $\langle \text{SV} \rangle \text{ MOD}=\textit{ind NB}=\textit{X PERS}=\textit{Y GR}=\textit{Z}$
 (2) $\langle \text{SN} \rangle \text{ NB}=\textit{X PERS}=\textit{Y GR}=\textit{Z}$ \rightarrow $\langle \text{DET} \rangle \text{ NB}=\textit{X PERS}=\textit{Y GR}=\textit{Z}$ $\langle \text{NC} \rangle \text{ NB}=\textit{X PERS}=\textit{Y GR}=\textit{Z}$
 (3) $\langle \text{SN} \rangle \text{ NB}=\textit{X PERS}=\textit{Y GR}=\textit{Z}$ \rightarrow $\langle \text{PRO} \rangle \text{ NB}=\textit{X PERS}=\textit{Y GR}=\textit{Z}$
 (4) $\langle \text{SV} \rangle \text{ MOD}=\textit{M NB}=\textit{X PERS}=\textit{Y}$ \rightarrow $\langle \text{V} \rangle \text{ MOD}=\textit{M NB}=\textit{X PERS}=\textit{Y SSCAT}=(\textit{ })$
 (5) $\langle \text{SV} \rangle \text{ MOD}=\textit{M NB}=\textit{X PERS}=\textit{Y}$ \rightarrow $\langle \text{V} \rangle \text{ MOD}=\textit{M NB}=\textit{X PERS}=\textit{Y SSCAT}=(\textit{GN})$ $\langle \text{GN} \rangle$
 (6) $\langle \text{SV} \rangle \text{ MOD}=\textit{M NB}=\textit{X PERS}=\textit{Y}$ \rightarrow $\langle \text{V} \rangle \text{ MOD}=\textit{M NB}=\textit{X PERS}=\textit{Y SSCAT}=(\textit{VCOMP})$ $\langle \text{SV} \rangle \text{ MOD}=\textit{inf}$
 (7) $\langle \text{SV} \rangle \text{ MOD}=\textit{M NB}=\textit{X PERS}=\textit{Y GR}=\textit{Z}$
 \rightarrow $\langle \text{V} \rangle \text{ MOD}=\textit{M NB}=\textit{X PERS}=\textit{Y SSCAT}=(\textit{ADJ})$ $\langle \text{ADJ} \rangle \text{ NB}=\textit{X PERS}=\textit{Y GR}=\textit{Z}$

1 — Donnez la représentation du lexique correspondant à cette grammaire pour le vocabulaire suivant : { il, le, la, bébé, soupe, crie, mange, manger, aller, est, veut, content }

2 — Donnez le char complet (agenda) correspondant à l'analyse descendante tabulaire des énoncés suivants :

a) *le bébé est content*

b) *il veut manger la soupe*

3 — Traduisez maintenant l'ensemble de cette grammaire (règles + lexique) sous la forme d'une grammaire LFG.

4 — Donnez la c-structure, la f-structure et la structure prédicative de l'énoncé : *le bébé veut aller manger*.

La suite de cet exercice est hors programme cette année...

5 — Traduisez maintenant cette grammaire dans le formalisme des grammaires d'arbres adjoints (TAG).

6 — Donnez la construction des dérivations (succession d'ajonctions, substitutions et unifications) correspondant à l'analyse de l'énoncé *le bébé crie*.

Exercice 4 : Grammaire lexicale-fonctionnelle

On considère la grammaire LFG ci-dessous :

- (1) $\langle S \rangle \rightarrow \langle \text{SN} \rangle \langle \text{SV} \rangle$
 $(\uparrow \text{SUJ}) = \downarrow \quad \uparrow = \downarrow$
 $(\uparrow \text{NB}) = (\downarrow \text{NB})$
 $(\uparrow \text{PERS}) = (\downarrow \text{PERS})$
- (2) $\langle \text{SN} \rangle \rightarrow \langle \text{NP} \rangle$
 $\uparrow = \downarrow$
- (3) $\langle \text{SN} \rangle \rightarrow \langle \text{PRO} \rangle$
 $\uparrow = \downarrow$
- (4) $\langle \text{SV} \rangle \rightarrow \langle \text{GV} \rangle \{ \langle \text{ADV} \rangle \}$
 $\uparrow = \downarrow \quad (\uparrow \text{ADJ}) \ni \downarrow$
- (5) $\langle \text{SV} \rangle \rightarrow \langle \text{GV} \rangle \{ \langle \text{ADV} \rangle \} \langle \text{COMP} \rangle$
 $\uparrow = \downarrow \quad (\uparrow \text{ADJ}) \ni \downarrow \quad (\uparrow \text{VCOMP}) = \downarrow$
 $(\downarrow \text{FOCUS}) = (\uparrow \text{SUJ})$
- (6) $\langle \text{GV} \rangle \rightarrow \{ \langle \text{CLITIQUE} \rangle \} \langle \text{V} \rangle \{ \langle \text{SN} \rangle \}$
 $(\uparrow \text{OBJ}) = \downarrow \quad \uparrow = \downarrow \quad (\uparrow \text{OBJ}) = \downarrow$
- (7) $\langle \text{COMP} \rangle \rightarrow \langle \text{GAP} \rangle \langle \text{SV} \rangle$
 $(\uparrow \text{SUJ}) = (\uparrow \text{FOCUS}) \quad \uparrow = \downarrow$
 $(\downarrow \text{FORM}) = \textit{inf}$

Lexique :

Jean : (NP, $(\uparrow \text{PRED}) = \textit{Jean}$, $(\uparrow \text{NB}) = \textit{sg}$, $(\uparrow \text{GR}) = \textit{m}$, $(\uparrow \text{PERS}) = 3$)

Lio : (NP, $(\uparrow \text{PRED}) = \textit{Jean}$, $(\uparrow \text{NB}) = \textit{sg}$, $(\uparrow \text{GR}) = \textit{f}$, $(\uparrow \text{PERS}) = 3$)

le : (CLITIQUE, $(\uparrow \text{PRED}) = \textit{le}$, $(\uparrow \text{NB}) = \textit{sg}$, $(\uparrow \text{GR}) = \textit{m}$, $(\uparrow \text{PERS}) = 3$)

cela : (PRO, $(\uparrow \text{PRED}) = \textit{cela}$, $(\uparrow \text{NB}) = \textit{sg}$, $(\uparrow \text{GR}) = \textit{m}$, $(\uparrow \text{PERS}) = 3$)

veut : (V, $(\uparrow \text{PRED}) = \textit{vouloir} \langle (\uparrow \text{SUJ}) (\uparrow \text{VCOMP}) \rangle$, $(\uparrow \text{NB}) = \textit{sg}$, $(\uparrow \text{PERS}) = 3$, $(\uparrow \text{FORM}) = \textit{ind_prt}$)

veut : (V, $(\uparrow \text{PRED}) = \textit{vouloir} \langle (\uparrow \text{SUJ}) (\uparrow \text{OBJ}) \rangle$, $(\uparrow \text{NB}) = \textit{sg}$, $(\uparrow \text{PERS}) = 3$, $(\uparrow \text{FORM}) = \textit{ind_prt}$)

∅ : (GAP)

donner (V, $(\uparrow \text{PRED}) = \textit{donner} \langle (\uparrow \text{SUJ}) (\uparrow \text{OBJ}) \rangle$, $(\uparrow \text{FORM}) = \textit{inf}$)

rapidement : (ADV $(\uparrow \text{PRED}) = \textit{rapidement}$)

- 1) Donnez la c-structure et la f-structure **complètes** de l'énoncé : *Jean le veut rapidement.*
- 2) Regardez la règle (6). Pourquoi la grammaire n'autorisera-t-elle pas l'énoncé *Jean le veut Lio* ?
- 3) Regardez les règles (5) et (7). A quoi sert la notion de Gap ? Et quel est le rôle de l'attribut FOCUS ?
- 4) Donnez la c-structure et la f-structure **complètes** correspondant à l'analyse de : *Jean veut donner cela*

Exercice 5

Les dates et horaires suivent une syntaxe bien particulière qui nécessite la construction de grammaires spécifiques pour les analyser. Dans cet exercice, on demande de créer une grammaire modélisant les horaires en français. Cette grammaire devra pouvoir générer les séquences suivantes :

<i>deux heures</i>	<i>deux heures dix</i>	<i>deux heures moins quinze</i>
<i>deux heures et demie</i>	<i>deux heures et quart</i>	<i>deux heures moins le quart</i>
<i>deux heures quarante trois</i>	<i>midi</i>	<i>midi quinze</i>

Elle devra évidemment ne pas surgénérer de séquences erronées comme par exemple :

<i>Deux heures moins demie</i>	<i>vingt-huit heures cinq</i>	<i>deux heures soixante douze</i>
--------------------------------	-------------------------------	-----------------------------------

1 — Ecrire une grammaire attribuée qui génère ce langage. Le pouvoir de génération fort de cette grammaire devra être tel qu'à la fin de l'analyse, l'arbre généré distinguera les heures des minutes.

2 — Pouvez-vous traduire cette grammaire dans le formalisme des LFG de telle manière que les heures et minutes soient deux attributs (ou arguments) des structures fonctionnelles et prédicatives finales.

Exercice 6 : gestion de clitique

On désire réaliser une grammaire qui traite correctement les clitiques en français. On se contentera de gérer ici les énoncés comprenant des sujets et compléments pronominaux combinés à un verbe éventuellement complété par un modal. Par exemple, notre grammaire devra générer les énoncés suivants :

<i>je le vois</i>	<i>je le lui donne</i>	<i>je lui parle</i>
<i>je veux le voir</i>	<i>je veux le lui donner</i>	<i>je veux lui parler</i>

Elle devra rejeter des énoncés présentant des clitiques mal positionnés ou erronés d'un point de vue morphologique :

<i>me lui parler</i>	<i>je parler lui</i>	<i>je lui le donne</i>
----------------------	----------------------	------------------------

Ecrire une grammaire attribuée qui génère ce langage, puis donner sa traduction en LFG.

Exercice 7 : Mots contractés et analyse syntaxique lexicalisée

Il existe en français un certain nombre de mots dits contractés qui correspondent fonctionnellement à l'association de plusieurs parties du discours. Par exemple, l'article contracté *au* correspond à l'association de la préposition *à* et du déterminant *le*. De même pour *du* = *de* + *le*.

1. Comment traiteriez-vous ce type de mot dans le cadre d'une grammaire syntagmatique classique, ou encore d'une grammaire lexicale-fonctionnelle ?
2. Pourquoi est-il possible de traiter cette spécificité d'une manière plus élégante avec une grammaire entièrement lexicalisée ?
3. Comment traiter précisément les mots contractés *au* et *du* avec une grammaire d'arbres adjoints (TAG) ?

Variabilité langagière et langage oral

Exercice 1 : Jeux de langues

Le langage dispose de ressources multiples pour exprimer une idée particulière. Considérons par exemple l'énoncé :

La boisson préférée de l'actuel président de la République est la bière Corona

1. Donnez quatre formulations différentes de cet énoncé qui, tout en reprenant les mêmes mots lexicaux, permettent d'insister sur un élément ou un autre de l'énoncé. A chaque fois, on soulignera l'élément mis en relief et on nommera le procédé langagier utilisé.
2. A quel procédé langagier répond l'insertion de la fonction « l'actuel président de la République » dans l'énoncé ci-dessous :

La boisson préférée de Jacques Chirac, l'actuel président de la République, est la bière Corona

Exercice 2 : Faits de langue

- 1) Sauriez-vous typer les différentes relatives de français populaire ci-dessous, qui correspondent toutes à des énoncés attestés :

- (a) *C'est une petite ville où il ferait assez bon y vivre*
- (b) *J'ai vendu ma petite maison que je tenais tant*
- (c) *C'est un copain que j'ai passé mon enfance avec lui*

- 2) On considère le dialogue page suivante, adapté du corpus Air France. Donnez l'ensemble des réparations présentes dans cet extrait en précisant à chaque fois leur type (répétition simple ou avec enrichissement lexical, corrections). On présentera les énoncés avec réparation sous la forme d'une analyse en grille.

C1 --	bonjour madame pourriez-vous me dire s'il existe des avions pour Catane en Sicile
O2 --	excusez-moi monsieur la ligne est mauvaise je n'ai pas entendu votre question
C2 --	est-ce qu'il existe des avions des vols pour Catane en Sicile
O3 --	en Sicile et vous souhaiteriez partir à quelle période de l'année
C3 --	euh vers le 16 mai
O4 --	je vous donne la fréquence des vols directs pour le mois euh à partir du mois de de mai alors justement nous aurons un vol direct qui partira tous les samedis
C4 --	samedi
O5 --	de de Charles oui de Charles de Gaulle à 18h25
C5 --	Charles de Gaulle 18h25
O6 --	et l'arrivée à Catane en vol direct se fera à 21 heures
C6 --	21 heures /--/ bon ça c'est bon /-/ et c'est valable à peu près pour le mois de mai
O7 --	c'est valable pour le mois de mai ce que je vous donne c'est pour le mois de mai
C7 --	c'est bien pour Catane
O8 --	tout à fait
C8 --	est-ce qu'il y a des des des des prix de vols vacances
O9 --	il y a un vol vacances je vous donne peut-être l'horaire du retour et la fréquence du retour
C9 --	oh le retour c'est pas grave enfin est-ce que vous pouvez me donner m'indiquer le prix
O10 --	oui bien sûr /-/ alors le retour est important les retours la fréquence c'est le dimanche

- 3) Donnez l'ensemble des extractions présentes dans cet extrait, en précisant à chaque fois leur type (inversion simple, dislocation, clivée ou pseudo-clivée).

Exercice 3 : Faits de langue

Si le français peut être considéré comme un langage à ordre relativement fixe, de nombreuses constructions orales ou écrites enfreignent cependant l'ordre canonique SVO (sujet-verbe-objets et autres compléments) qu'est sensée suivre cette langue. On parle souvent d'extraction pour caractériser ces constructions qui modifient l'ordre standard des mots dans l'énoncé : un élément de l'énoncé a été « extrait » de sa place habituelle. Ces extractions sont réalisées suivant quatre types principaux de procédés (Gadet, 1992) :

— les **inversions**, qui correspondent à une modification de la position d'un groupe ne se manifestant par aucune autre marque linguistique. Par exemple :

- (1a) *sur Héraklion on n'a qu'un seul tarif spécial*
- (1b) *jusqu'à 21 h 30 euh vous pouvez manger au restaurant*

— les **dislocations**, que l'on appelle également **doubles-marquages** (Blanche-Benveniste, 1997). Dans ce cas, le déplacement est marqué par un clitique de reprise rappelant le groupe ainsi que sa fonction. Dans l'exemple qui suit, l'argument *le visa* est ainsi repris par le clitique *le* :

(2a) *le visa on l'a au consulat*

Cette reprise peut également être réalisée à l'aide de l'expression *c'est* ou plus généralement *çà + <verbe>* :

(2b) *donc l'office du tourisme c'est c'est à côté du stade*

— les **présentatifs**, dans lesquels un élément initial (*c'est* ou les introductifs construits avec le verbe *avoir* : *il y a / j'ai / on a* etc.) introduit explicitement la partie de l'énoncé détachée et est suivi d'une (fausse) subordonnée introduite par *qui* ou *que*. Par exemple :

(3a) *j'ai quelqu'un qu'est allé prendre des billets charters pour moi*

(3b) *c'est bien le renseignement que vous vouliez avoir*

Le procédé en *c'est ... que* est qualifié de structure clivée. On peut également caractériser une structure pseudo-clivée qui réalise une double extraction suivant le motif *ce que ... c'est* :

(3c) *ce que vous vouliez avoir c'est bien ce renseignement*

— Enfin, les énoncés **binaires** se caractérisent par un élément détaché qui perd toute dépendance claire avec le reste de l'énoncé, souvent du fait d'une ellipse plus ou moins évidente à détecter. Par exemple :

(4) *le tarif vacances euh toute modification entraîne des frais*

On donne ci-dessous un extrait du corpus oral *Air France* réalisé par l'équipe de Marie-Annick Morel à l'Université de Paris – Sorbonne Nouvelle.

1 — Caractériser au sein de cet extrait de corpus différents énoncés correspondant à cette typologie de procédés. On cherchera ainsi à trouver un exemple d'extraction pour chacun des procédés suivants :

- Inversion,
- Double marquage par pronom ; double marquage par *c'est*,
- Présentatif simple, clivé et pseudo-clivé,
- Énoncé binaire

2 — L'exemple ci-dessous est également extrait du corpus *Air France* :

(5) *sur le Nord c'est par Air France qu'on peut réserver*

Pouvez-vous montrer en quoi cet énoncé disloqué pose problème à une analyse syntagmatique classique.

COMMUNICATION I-49 Erreur ! Signet non défini.

O1 --	Air France bonjour
C1 --	<i>bonjour madame pourriez-vous me dire s'il existe des avions pour Catane en Sicile</i>
O2 --	excusez-moi monsieur la ligne est mauvaise je n'ai pas entendu votre question
C2 --	<i>est-ce qu'il existe des avions, des vols pour Catane en Sicile</i>
O3 --	en Sicile et vous souhaiteriez partir à quelle période de l'année
C3 --	<i>e: vers le 16 mai, par là</i>
O4 --	/-/ h: allô alors je vous donne la fréquence des vols directs un, pou e: à partir du mois de de mai justement nous aurons un vol direct qui partira tous les samedis
C4 --	<i>samedi</i>
O5 --	de de ch, oui de Charles de Gaulle à 18h25
C5 --	<i>Charles de Gaulle 18h25</i>
O6 --	et l'arrivée à Catane, en vol direct /-/ se fera à 21 heures
C6 --	<i>21 heures /-/ bon ça c'est bon /-/ que valable à peu près pour le mois de mai</i>
O7 --	c'est valable pour le mois de mai ce que je vous donne c'est pour le mois de mai
C7 --	<i>c'est bien pour Catane</i>
O8 --	tout à fait
C8 --	<i>est-ce qu'il y a des des des des prix de vols vacances</i>
O9 --	dans la mesure où ce vol direct existe maintenant il y a un vol vacances je vous donne peut-être l'horaire du retour et la fréquence du retour
C9 --	<i>oh le retour c'est pas grave enfin est-ce que vous pouvez me don: m'indiquer le prix</i>

O10 --	oui bien sûr /-/ alors le retour est important les retours la fréquence c'est le dimanche un donc ça vous aidera à construire votre voyage
C10 --	<i>bon retour dimanche alors l'heure l'heure si vous pouvez me dire</i>
O11 --	départ 8h40
C11 --	<i>retour 8H40</i>
O12 --	et l'arrivée à Paris se fait à 11h30
C12 --	<i>oh la oh la la il faut le prendre de bonne heure oh la la oui bon</i>
O13 --	alors du point de vue tarif /-/ le tarif le plus intéressant donc le tarif visite minimum un week-end sur place maximum trois mois
C13 --	<i>oui</i>
O14 --	et il fait 3570 francs
C14 --	<i>3000:</i>
O15 --	3570 francs
C15 --	<i>3570 francs est-ce qu'il y a des conditions pour les personnes âgées de plus de 65 ans</i>
O16 --	non, il n'y a pas de conditions particulières le tarif vacances c'est le moins cher
C16 --	<i>ah le tarif vacances c'est le moins cher pour les personnes âgées y a rien</i>
O17 --	voilà
C17 --	<i>/-/ pour les couples</i>
O18 --	non plus non plus y aurait éventuellement une réduction mais qui serait calculée sur le plein tarif /-/ qui est beaucoup plus cher et ça: ce serait moins intéressant que deux vols vacances um
C18 --	<i>d'accord merci beaucoup</i>
O19 --	je veux
C19 --	<i>on est bien d'accord c'est bien pour Catane en Sicile</i>
O20 --	Catane en Sicile monsieur tout à fait
C20 --	<i>merci beaucoup</i>
O21 --	je
C21 --	<i>merci</i>
O22 --	je vous en prie

Exercice 4 : Coordination et réparation

Les répétitions et autres réparations se traduisent par un entassement paradigmatique qui consiste le plus souvent à reprendre le même syntagme jusqu'à arriver à la dénomination recherchée. Ce que traduit bien l'analyse en grille de C. Blanche-Benveniste :

*A chaque fois que je regarde **une partie**
un match
un match de tennis sur terre battue je m'ennuie.*

Du fait de cette pause syntagmatique, les différents éléments de la réparation remplissent toujours la même fonction syntaxique dans l'énoncé. Dans l'exemple ci-dessus, la réparation concerne ainsi la répétition du complément d'objet du verbe *regarde*. On se retrouve donc dans un schéma proche de celui des coordinations.

- 4) En partant de cette constatation, sauriez-vous donner une structure syntaxique possible de l'énoncé suivant :

il détestait les gens de Vichy enfin les collaborateurs du régime de Vichy

Les énumérations ne sont qu'une forme de coordination particulière où les premiers éléments coordonnés ne sont pas marqués par une conjonction de coordination, comme le montrent les exemples ci-dessous :

Coordination : *mes couleurs préférées sont le jaune et le bleu*

Enumération : *mes couleurs préférées sont le jaune le vert et le bleu*

- 5) En partant de cette constatation, sauriez-vous donner une structure syntaxique de l'énoncé suivant :

les trois constructeurs français d'automobiles sont Citroën Peugeot et Renault

- 6) Pouvez-vous en déduire une structure syntaxique possible de l'énoncé suivant :

on combattait les Vichysois les gens de Vichy enfin les collaborateurs avec l'occupant allemand.

Exercice 5 : Extractions, analyse syntagmatique ou en dépendances

Un des arguments centraux de Lucien Tesnière dans sa critique de l'approche syntagmatique était la difficulté des théories Chomskyennes à rendre compte des phénomènes d'extractions. Considérons l'inversion :

Dès son arrivée il fut accueilli chaleureusement

- 1) Est-on en présence d'une extraction projective ou non (i.e. discontinue) ?
- 2) L'élément extrait *dès son arrivée* est un complément du groupe verbal. Que peut-on dire alors de la structure syntaxique correspondant à l'énoncé dans une approche purement syntagmatique ?
- 3) Donnez alors la représentation syntaxique de l'énoncé dans une approche relevant des grammaires de dépendances (grammaires de liens). Le (pseudo...) problème observé dans la question précédente est-il encore présent ?