

UNIVERSITE PARIS OUEST NANTERRE LA DEFENSE

ECOLE DOCTORALE 139 : CONNAISSANCE, LANGAGE, MODELISATION

LABORATOIRE MoDYCo (UMR 7114)

La liaison en français : descriptions et analyses dans le corpus PFC

Thèse de doctorat nouveau régime

présentée

par

Géraldine-M. MALLET

Discipline : Sciences du Langage

sous la direction de Bernard LAKS

Composition du jury :

- Mme Marie-Hélène COTE, Professeure agrégée, Université d'Ottawa
- M. Pierre ENCREVE, Professeur et Directeur d'Etudes émérite, EHESS
- M. Marc KLEIN, Maître de Conférences, Université Paris Ouest Nanterre La Défense
- M. Bernard LAKS, Professeur, Université Paris Ouest Nanterre La Défense

Remerciements

Je tiens à remercier tout d'abord mon directeur de recherche, Bernard Laks, qui m'a encouragée, guidée et a fait preuve d'une grande patience à mon égard. Je lui suis également très reconnaissante d'avoir toujours fait en sorte que je puisse travailler dans les meilleures conditions qui soient.

Mes remerciements vont également à Marie-Hélène Côté et Pierre Encrevé qui ont accepté la tâche de rapporteurs, ainsi qu'à Marc Klein qui a bien voulu faire partie de mon jury.

Au sein du projet *Phonologie du Français Contemporain (PFC)*, je remercie tout particulièrement Jacques Durand et Chantal Lyche pour leurs encouragements et leur bienveillance. PFC a également été un « lieu » de rencontre et d'échange fructueux. Merci donc à Julien Eychenne, Gudrun Ledegen, Elissa Sobotta et Gabor Turcsan pour leur enthousiasme et leur disponibilité.

J'ai eu beaucoup de plaisir à être doctorante à MoDyCo et ma reconnaissance s'adresse aux membres du laboratoire, notamment à Marc Klein pour son soutien et les discussions que nous avons eues. Merci à René Lavie, Philippe Gréa, Sylvain Loiseau, Atanas Tchobanov et Ali Tifrit pour leurs encouragements et leur aide. Merci à Xavier Monnin et à Sophie de Pontonx d'avoir répondu présents dans l'urgence et enfin, merci à l'ensemble des membres du « tea party club ». Merci à Danielle Joseph pour sa gentillesse, sa bonne humeur constante et tout ce qu'elle a fait pour me « faciliter » la vie administrative.

Il est des gens sans lesquels ce travail ne serait pas. Aussi tout particulièrement, je tiens à remercier Sophie David pour le soutien qu'elle m'a apporté, pour toutes ces discussions qui m'ont permis d'avancer et d'accepter. Oui Sophie, sans tes encouragements, sans ton aide et sans ton amitié tout au long de ces années, cette thèse ne serait simplement pas.

Je tiens également à remercier mes parents pour le soutien qu'ils m'ont apporté, pour la confiance sans faille dont ils ont toujours fait preuve et qui m'a si souvent fait défaut. A vous je dois tout et la thèse n'est qu'un détail. Aude, ma courageuse, merci de reprendre le flambeau. Ca m'a fait comprendre que chacun a droit, à son tour, à

l'attention du clan au complet et qu'il y a un temps pour commencer mais aussi pour terminer. Merci pour les deux merveilles !

Une mention spéciale est décernée à François pour en avoir, stoïquement, autant supporté sans comprendre exactement pourquoi certaines phases pouvaient être si compliquées.

Merci pour leur amitié et pour tous les moments partagés : Sophie, Florence, Linda (oui, oui, merci pour la nuit blanche passée à tout relire), Claudie et Bertrand qui va devoir se faire une raison : non, ma thèse n'est pas sur le « e muet ».

Enfin, merci au « R. I. Kitz »: grâce à votre gentillesse et à votre générosité, ces derniers mois ont presque été heureux. Je n'oublierai pas.

Table des matières

REMERCIEMENTS.....	3
TABLE DES MATIERES.....	5
INDEX FIGURES ET TABLEAUX.....	9
CHAPITRE 1 INTRODUCTION	13
PREMIERE PARTIE : HISTOIRE DES TRAITEMENTS DE LA LIAISON.....	19
CHAPITRE 2 PARCOURS HISTORICO-CONCEPTUEL DE LA LIAISON	21
2.1 RAPPEL HISTORIQUE	22
2.2 LES NORMATIVISTES.....	24
2.3 LES PRESCRIPTIVISTES	27
2.4 LES DESCRIPTIVISTES	30
2.5 BILAN DU PARCOURS HISTORICO-CONCEPTUEL	34
CHAPITRE 3 LES DIMENSIONS D'ANALYSE.....	37
3.1 PHONOLOGIQUE.....	38
3.1.1 PRINCIPES GENERAUX DU PHENOMENE.....	38
3.1.2 UNE CONSONNE DE LIAISON	39
3.1.2.1 Bref inventaire	40
3.1.2.2 Des changements phonétiques.....	42
3.1.2.3 Synthèse.....	43
3.1.3 UNE STRUCTURE SYLLABIQUE	43
3.1.3.1 Représentation formelle de la syllabe.....	44
3.1.3.2 Un enchaînement ou pas.....	45
3.1.4 LE HIATUS.....	47
3.1.5 LES CARENCES D'UNE EXPLICATION PHONOLOGIQUE.....	54
3.2 MORPHO-SYNTAXIQUE.....	57
3.2.1 LA LIAISON EN /z/ : LE PLURIEL.....	59
3.2.2 LES ANALYSES PAR SUITE LINEAIRE DE CATEGORIES	61
3.2.3 PARTITION DES LIAISONS	64
3.2.3.1 Terminologie	65
3.2.3.2 Le contenu des classes de liaison selon Delattre (1966) et leur révision.....	67
3.2.3.2.1 La classe des erratiques.....	69
3.2.3.2.2 La classe des catégoriques	70
3.2.3.2.3 La classe des variables	75

Table des matières

3.2.3.2.4	Bilan.....	76
3.3	SOCIOLINGUISTIQUE.....	80
3.3.1	L'ARCHITECTURE VARIATIONNELLE.....	81
3.3.1.1	La variation inter-locuteurs.....	82
3.3.1.2	La variation intra-locuteurs.....	83
3.3.2	LA VARIATION SOCIOLINGUISTIQUE ET LA TRIPARTITION.....	84
3.3.2.1	Absence de variation : catégoriques et erratiques.....	84
3.3.2.2	Les variables comme lieu d'observation de la variation.....	85
3.3.3	LES FACTEURS SOCIOLOGIQUES DE LA VARIATION DE LA LIAISON.....	87
3.3.3.1	Le facteur stylistique.....	88
3.3.3.2	Le profil socioculturel.....	90
3.3.3.3	L'âge et le sexe.....	92
3.3.4	LES PRODUCTIONS AGRAMMATICALES.....	93
3.3.4.1	Dans les contextes de liaison catégorique.....	93
3.3.4.2	Dans les contextes de liaison erratique.....	93
3.3.5	BILAN.....	95
3.4	LEXICALE.....	96
3.4.1	LA LONGUEUR.....	97
3.4.1.1	La longueur de M1.....	98
3.4.1.2	La longueur du groupe syntaxique auquel appartient M2.....	100
3.4.1.3	La longueur du groupe syntaxique auquel appartient M1.....	102
3.4.1.4	Bilan.....	102
3.4.2	LA FREQUENCE.....	103
3.4.2.1	Fréquence de M1 et de M2.....	104
3.4.2.2	Fréquence de la co-occurrence.....	105
3.5	BILAN SUR LES DIMENSIONS D'ANALYSE.....	106
CHAPITRE 4	UN PROCESSUS PHONOLOGIQUE.....	107
4.1	L'APPROCHE GENERATIVE LINEAIRE ET LES REGLES.....	108
4.1.1	LA TRONCATION.....	108
4.1.2	LA METATHESE.....	114
4.1.3	L'EPENTHESE.....	117
4.1.4	BILAN D'UNE APPROCHE PAR REGLES.....	120
4.2	L'APPROCHE MULTILINEAIRE ET LES REPRESENTATIONS.....	121
4.3	LA THEORIE DE L'OPTIMALITE ET LA NOTION DE CONTRAINTE.....	128
4.4	THEORIES BASEES SUR L'USAGE.....	130
4.5	LE STATUT LEXICAL DE LA CONSONNE DE LIAISON.....	136
4.5.1	LA CL EST RATTACHEE A M1.....	137
4.5.2	LA CL EST RATTACHEE A M2.....	139
4.5.3	LA CL EST EPENTHETIQUE.....	140
4.5.4	LA CONSONNE APPARTIENT A LA CONSTRUCTION.....	141
4.5.5	BILAN.....	142
SECONDE PARTIE : ANALYSES EMPIRIQUES.....		143
CHAPITRE 5	METHODOLOGIE.....	145
5.1	LE PROJET PFC.....	145
5.1.1	PROTOCOLE D'ENQUETE.....	147

5.1.1.1	L'entretien guidé	150
5.1.1.2	L'entretien libre.....	150
5.1.1.3	La liste de mots.....	151
5.1.1.4	Le texte	151
5.1.2	TRAITEMENT DES ENQUETES	152
5.1.2.1	Principes techniques	152
5.1.2.2	Système de transcription adopté	154
5.1.2.3	Conventions de transcription	156
5.1.2.4	La transcription orthographique et ses limites.....	159
5.1.3	LE CODAGE LIAISON	160
5.1.3.1	Principes généraux.....	160
5.1.3.2	Codage alphanumérique	161
5.1.3.3	Les limites du codage	163
5.1.3.3.1	Le transcrip-teur codeur qui fait des erreurs	163
5.1.3.3.2	Le manque d'informations	163
5.1.3.3.3	Le problème de la granularité du codage	164
5.1.3.3.4	Le problème de la complétude des contextes.....	164
5.1.3.3.5	Les « choix théoriques implicites » opérés par le codeur	165
5.1.3.3.6	Bilan : des choix méthodologiques dans la dépendance des choix théoriques.....	165
5.2	CONSTITUTION ET DESCRIPTION DU CORPUS.....	166
5.2.1	LES POINTS D'ENQUETE	166
5.2.1.1	Couverture géographique	168
5.2.1.2	Etat d'avancement	169
5.2.2	LES LOCUTEURS	170
5.2.2.1	Répartition géographique	170
5.2.2.2	Sexe et âge des locuteurs	172
CHAPITRE 6	METHODOLOGIE, RESULTATS ET ANALYSES	175
6.1	INTRODUCTION	176
6.1.1	LES CONTEXTES DE LIAISON NON CONSIDERES	179
6.1.2	LES CONTEXTES DE LIAISON CONSIDERES.....	185
6.2	LA LIAISON ET LES FACTEURS SOCIOLINGUISTIQUES	188
6.2.1	FREQUENCE DE LA LIAISON EN FONCTION DU STYLE	188
6.2.2	FREQUENCE DE LA LIAISON EN FONCTION DU NIVEAU D'ETUDES.....	190
6.2.3	FREQUENCE DES LIAISONS EN FONCTION DE LA CLASSE D'AGE.....	193
6.2.4	BILAN	195
6.2.4.1	Variation diaphasique	195
6.2.4.2	Variation diastratique	201
6.3	LES CONSONNES DE LIAISON	204
6.3.1	INVENTAIRE ET DISTRIBUTION DES CONSONNES DE LIAISON	210
6.3.2	FREQUENCE DE CHAQUE CL DANS CHAQUE ECL	212
6.3.3	ANALYSE CONTEXTUELLE D'APPARITION DES CONSONNES DE LIAISON	214
6.3.3.1	Les consonnes de liaison non considérées.....	214
6.3.3.1.1	La consonne /r/.....	215
6.3.3.1.2	La consonne /p/	215
6.3.3.1.3	La consonne /k/	217
6.3.3.2	Les consonnes de liaison considérées.....	218
6.3.3.2.1	La consonne /n/	219
6.3.3.2.2	La consonne /t/	227
6.3.3.2.3	La consonne /z/	240
6.3.3.3	Bilan	255
6.4	INFLUENCE DE LA LONGUEUR	260

Table des matières

6.4.1	REPARTITION DES MONOSYLLABES ET DES POLYSYLLABES	262
6.4.2	LA LIAISON AVEC LES MONOSYLLABES.....	266
6.4.3	LA LIAISON AVEC LES POLYSYLLABES	268
6.4.4	BILAN	270
6.5	SYNTAXE.....	271
6.5.1	CONJONCTION MONOSYLLABIQUE +	276
6.5.2	PREPOSITION MONOSYLLABIQUE +	277
6.5.3	PREPOSITION POLYSYLLABIQUE +	279
6.5.4	ADVERBE MONOSYLLABIQUE +	280
6.5.5	LES TEMPS SIMPLES DE <i>ETRE</i> + ET <i>AVOIR</i> +	282
6.5.5.1	<i>être</i> +	282
6.5.5.2	<i>avoir</i> +	284
6.5.6	BILAN	286
CHAPITRE 7	CONCLUSION GENERALE.....	289
	REFERENCES BIBLIOGRAPHIQUES	297
	LISTE DES ANNEXES	311
Annexe 1	Texte	311
Annexe 2	Liste de mots.....	312
Annexe 3	Fiche signalétique	313
Annexe 4	Consentement de participation	315
Annexe 5	Répartition des locuteurs par secteur et par sexe.....	316
Annexe 6	Répartition des locuteurs par tranche d'âge.....	317
Annexe 7	Occurrences de liaisons codées « incertaines »	318
Annexe 8	Contextes de liaison identifiés dans le texte PFC.....	319
Annexe 9	Liste des formes identifiées	320
Annexe 10	Contextes de liaisons en /z/ identifiés.....	322
Annexe 11	Contextes de liaisons en /n/ identifiés	323
Annexe 12	Contextes de liaisons en /t/ identifiés	324
Annexe 13	Ensemble des contextes de liaisons erratiques	325
Annexe 14	Contextes de liaison où les deux catégories (M1 et M2) sont spécifiées	326
Annexe 15	Ensemble des contextes de liaison dans lesquels seul M2 est spécifié.....	327
Annexe 16	Ensemble des contextes de liaison faisant partie de classes ouvertes.....	328

Index Figures et Tableaux

Figure 3-1	Synthèse sur les caractéristiques de chaque CL.....	43
Figure 3-2	Structure syllabique minimale	44
Figure 3-3	Représentation où la consonne flottante appartient au premier mot (Encrevé 1988 : 172).....	46
Figure 3-4	Association du segment flottant à l'attaque flottante qui le suit (Encrevé, 1988 : 177)	46
Figure 3-5	La liaison et le hiatus (Ågren, 1973).....	52
Figure 3-6	Tableau simplifié de la classification des liaisons (Delattre, 1947 : 152).....	68
Figure 3-7	Tableau de la classe des liaisons erratiques	70
Figure 3-8	Tableau des contextes de liaisons catégoriques	71
Figure 3-9	Réalisation de la liaison dans la construction « adverbe <i>pas</i> suivi d'un M2 » dans les débats de 1974 et 1981 entre VGE et FM (Encrevé, 1988 : 64 - 65).....	74
Figure 3-10	Tableau des contextes « catégoriques » intégrant tout ou partie la classe des liaisons variables.....	75
Figure 3-11	Tableau des contextes de liaisons variables	76
Figure 3-12	Nouvelle répartition des contextes de liaisons erratiques	78
Figure 3-13	Nouvelle répartition des contextes de liaisons catégoriques.....	79
Figure 3-14	Nouvelle répartition des contextes de liaisons variables.....	80
Figure 3-15	Représentation de la variation (Gadet, 2007 : 23).....	82
Figure 3-16	Dimensions et facteurs de variation de la liaison.....	88
Figure 3-17	Identification des registres dans la représentation traditionnelle, chez Fouché (1959) et chez Delattre (1966).....	90
Figure 4-1	Alternances Consonne ~ Zéro.....	107
Figure 4-2	French Truncation Rule (Schane, 1968 : 10)	109
Figure 4-3	Répartition des segments en classes naturelles	110
Figure 4-4	Chute de la voyelle	110
Figure 4-5	Chute de la consonne	110
Figure 4-6	Application de la French Truncation Rule.....	112
Figure 4-7	Règle d'effacement de la consonne finale.....	112
Figure 4-8	Règle unique de suppression de la consonne	113
Figure 4-9	Règle de Liaison	114
Figure 4-10	Règle de Troncation.....	115
Figure 4-11	Application des Règles de Liaison et de Troncation.....	115
Figure 4-12	Règle d'épenthèse.....	117
Figure 4-13	Règle d'insertion de <i>n</i>	118
Figure 4-14	Règle de dénasalisation de la voyelle	118
Figure 4-15	Règle d'insertion de <i>z</i>	118
Figure 4-16	Convention universelle d'association des codas flottantes	125
Figure 4-17	Représentation du (non)-rattachement des consonnes finales.....	126
Figure 5-1	Répartition des corpus de PFC en fonction des registres	148
Figure 5-2	Exemple de codes des points d'enquêtes et des locuteurs	150
Figure 5-3	Exemple de Textgrid.....	153
Figure 5-4	Exemple de transcription / codage dans Praat.....	154
Figure 5-5	Exemple de transcription « transpraatée »	157
Figure 5-6	Exemple de codage liaison	162
Figure 5-7	Points d'enquêtes PFC.....	167
Figure 5-8	Répartition géographique des enquêtes.....	168
Figure 5-9	Représentation graphique de la répartition géographique des enquêtes.....	169
Figure 5-10	Représentation graphique de l'état d'avancement des enquêtes PFC.....	169
Figure 5-11	Répartition des locuteurs par enquête	171
Figure 5-12	Représentation graphique de la répartition des enquêtés par secteur géographique	172
Figure 5-13	Répartition des locuteurs par tranche d'âge.....	172
Figure 5-14	Représentation graphique de la répartition des témoins par tranche d'âge.....	173
Figure 6-1	Corpus d'enquêtes utilisées	177
Figure 6-2	Nombre total de contextes « liaison » codés dans le corpus fourni par PFC	178

Index Figures et Tableaux

Figure 6-3	Liaisons non considérées	179
Figure 6-4	Contextes de liaison considérés	186
Figure 6-5	Réalisation de la liaison en lecture et en conversation.....	187
Figure 6-6	Réalisation de la liaison selon le registre	189
Figure 6-7	Représentation graphique de la réalisation de la liaison selon le registre	189
Figure 6-8	Réalisation de la liaison selon le niveau d'études	191
Figure 6-9	Représentation graphique de la réalisation de la liaison selon le niveau d'études	192
Figure 6-10	Réalisation de la liaison selon la classe d'âge.....	194
Figure 6-11	Représentation graphique de la réalisation de la liaison selon la classe d'âge.....	194
Figure 6-12	Répartition des corpus PFC selon les registres	199
Figure 6-13	Les propositions de distribution des CL /n/, /t/ et /z/	205
Figure 6-14	Ensemble des corpus d'observation créés pour chaque CL	209
Figure 6-15	Tableau des abréviations utilisées pour l'analyse des CL.....	210
Figure 6-16	Distribution des CL dans le corpus comprenant les liaisons réalisées	212
Figure 6-17	Fréquence des liaisons réalisées et non réalisées dans chaque ECL.....	213
Figure 6-18	Fréquence des liaisons réalisées et non réalisées dans chaque ECL sous forme graphique	213
Figure 6-19	Liaisons réalisées et non réalisées dans les ECL non considérés (<i>Er</i> , <i>Ep</i> et <i>Ek</i>).....	215
Figure 6-20	Fréquence des liaisons réalisées et non réalisées dans <i>Er</i>	215
Figure 6-21	Fréquence des liaisons réalisées et non réalisées dans <i>Ep</i>	216
Figure 6-22	Fréquence des liaisons réalisées et non réalisées dans <i>Ek</i>	217
Figure 6-23	Consonnes de liaison considérées	218
Figure 6-24	Liaisons réalisées et non réalisées dans les ECL considérées (<i>En</i> , <i>Et</i> et <i>Ez</i>).....	218
Figure 6-25	Tableau des abréviations utilisées pour l'analyse contextuelle des consonnes /n/, /t/ et /z/	219
Figure 6-26	Contextes identifiés pour les CL /z/, /n/ et /t/	219
Figure 6-27	Fréquence des liaisons réalisées et non réalisées dans <i>En</i>	220
Figure 6-28	Contextes identifiés pour la CL /n/ (<i>En</i>)	220
Figure 6-29	Fréquence des liaisons réalisées et non réalisées dans <i>En</i>	220
Figure 6-30	M1 contenus dans <i>En</i>	221
Figure 6-31	Fréquence des liaisons en /n/ avec <i>on</i>	222
Figure 6-32	Représentation de <i>on</i> dans <i>En</i> , EILRn et EILNRn	222
Figure 6-33	Fréquence des liaisons en /n/ avec <i>mon</i> et <i>son</i>	222
Figure 6-34	Représentation de <i>mon</i> dans <i>En</i> , EILRn et EILNRn	223
Figure 6-35	Représentation de <i>son</i> dans <i>En</i> , EILRn et EILNRn.....	223
Figure 6-36	Fréquence des liaisons en /n/ avec <i>un</i>	223
Figure 6-37	Représentation de <i>un</i> dans <i>En</i> , EILRn et EILNRn	224
Figure 6-38	Fréquence des liaisons en /n/ avec <i>en</i>	224
Figure 6-39	Fréquence de <i>en</i> dans <i>En</i> , EILRn et EILNRn.....	225
Figure 6-40	Fréquence des liaisons en /n/ avec <i>bien</i>	225
Figure 6-41	Représentation de <i>bien</i> dans <i>En</i> , EILRn et EILNRn	225
Figure 6-42	Fréquence de réalisation de la liaison des mots dans <i>En</i>	226
Figure 6-43	Représentation graphique des principales unités constitutives de EILRn et EILNRn.....	227
Figure 6-44	Fréquence des liaisons réalisées et non réalisées dans <i>Et</i>	228
Figure 6-45	Contextes identifiés pour la CL /t/ (<i>Et</i>).....	228
Figure 6-46	Fréquence des liaisons réalisées et non réalisées dans <i>Et</i>	228
Figure 6-47	M1 contenus dans <i>Et</i>	229
Figure 6-48	Fréquence des liaisons en /t/ avec les formes fléchies du verbe <i>être</i>	231
Figure 6-49	Représentation des formes fléchies de <i>être</i> dans <i>Et</i> , EILRt et EILNRt	232
Figure 6-50	Fréquence des liaisons en /t/ avec les formes fléchies du verbe <i>avoir</i>	232
Figure 6-51	Représentation des formes fléchies de <i>avoir</i> dans <i>Et</i> , EILRt et EILNRt.....	233
Figure 6-52	Fréquence des liaisons en /t/ avec les autres formes verbales.....	234
Figure 6-53	Représentation des autres formes verbales dans <i>Et</i> , EILRt et EILNRt.....	234
Figure 6-54	Fréquence des liaisons en /t/ avec <i>petit</i> et <i>grand</i>	235
Figure 6-55	Représentation de <i>petit</i> dans <i>Et</i> , EILRt et EILNRt	235
Figure 6-56	Représentation de <i>grand</i> dans <i>Et</i> , EILRt et EILNRt	236
Figure 6-57	Fréquence des liaisons en /t/ avec <i>tout</i>	236
Figure 6-58	Représentation de <i>tout</i> dans <i>Et</i> , EILRt et EILNRt.....	237
Figure 6-59	Fréquence des liaisons en /t/ avec <i>quand</i>	237

Figure 6-60	Représentation de <i>quand</i> dans <i>EIt</i> , <i>EILRt</i> et <i>EILNRt</i>	238
Figure 6-61	Fréquence de réalisation de la liaison dans <i>EIt</i>	238
Figure 6-62	Représentation graphique des principales unités constituantes de <i>EILRt</i> et <i>EILNRt</i>	239
Figure 6-63	Fréquence des liaisons réalisées et non réalisées dans <i>Ez</i>	241
Figure 6-64	Contextes identifiés pour la CL /z/ (<i>Ez</i>)	241
Figure 6-65	Fréquence des liaisons réalisées et non réalisées dans <i>EIz</i>	242
Figure 6-66	M1 contenus dans <i>EIz</i>	243
Figure 6-67	Fréquence des liaisons en /z/ avec les déterminants	244
Figure 6-68	Représentation des déterminants dans <i>EIz</i> , <i>EILRz</i> et <i>EILNRz</i>	245
Figure 6-69	Fréquence des liaisons en /z/ avec <i>deux</i> et <i>trois</i>	246
Figure 6-70	Représentation de <i>deux</i> et <i>trois</i> dans <i>EIz</i> , <i>EILRz</i> et <i>EILNRz</i>	246
Figure 6-71	Fréquence des liaisons en /z/ avec les pronoms	247
Figure 6-72	Représentation des pronoms dans <i>EIz</i> , <i>EILRz</i> et <i>EILNRz</i>	248
Figure 6-73	Fréquence des liaisons en /z/ avec les formes fléchies de <i>être</i>	249
Figure 6-74	Représentation des formes fléchies de <i>être</i> dans <i>EIz</i> , <i>EILRz</i> et <i>EILNRz</i>	249
Figure 6-75	Fréquence des liaisons en /z/ avec les autres formes verbales	250
Figure 6-76	Représentation des autres formes verbales dans <i>EIz</i> , <i>EILRz</i> et <i>EILNRz</i>	250
Figure 6-77	Fréquence des liaisons en /z/ avec les mots invariables	251
Figure 6-78	Représentation des mots invariables dans <i>EIz</i> , <i>EILRz</i> et <i>EILNRz</i>	252
Figure 6-79	Fréquence des liaisons en /z/ avec les autres mots	253
Figure 6-80	Représentation des autres mots dans <i>EIz</i> , <i>EILRz</i> et <i>EILNRz</i>	254
Figure 6-81	Fréquence de réalisation de la liaison des mots dans <i>EIz</i>	254
Figure 6-82	Représentation graphique des principales unités constituantes de <i>EILRz</i> et <i>EILNRz</i>	255
Figure 6-83	Représentation graphique de la réalisation de la liaison des CL /n/, /t/ et /z/ dans <i>EI</i>	256
Figure 6-84	Distribution des consonnes /n/, /z/ et /t/	256
Figure 6-85	Formes polysyllabique et monosyllabique selon la (non)-réalisation du schwa	258
Figure 6-86	Exemples de formes ayant le même nombre de contextes mais des taux de LR très distincts	259
Figure 6-87	Nombre de contextes pour les CL /z/, /n/ et /t/	262
Figure 6-88	Répartition des contextes de liaison en /n/, /t/ et /z/ selon la longueur	262
Figure 6-89	Répartition des formes monosyllabiques et polysyllabiques avec la CL /t/	263
Figure 6-90	Formes monosyllabiques ou polysyllabiques à finale /z/	264
Figure 6-91	Répartition des formes monosyllabiques et polysyllabiques avec la CL /z/	265
Figure 6-92	La liaison avec les monosyllabes	266
Figure 6-93	Les formes monosyllabiques liaisonnant en /t/	266
Figure 6-94	Les formes monosyllabiques liaisonnant en /z/	267
Figure 6-95	La liaison avec les polysyllabes	268
Figure 6-96	Les formes polysyllabiques liaisonnant en /t/	269
Figure 6-97	Les formes polysyllabiques liaisonnant en /z/	270
Figure 6-98	Tableau révisé de la grille tripartite de Delattre	272
Figure 6-99	Contextes de liaison considérés dans l'analyse syntaxique	275
Figure 6-100	Taux de LR en corpus des conjonctions monosyllabiques considérées	276
Figure 6-101	Taux de LR en corpus des prépositions monosyllabiques considérées	277
Figure 6-102	Taux de LR en corpus de la forme <i>en</i>	278
Figure 6-103	Taux de LR en corpus des prépositions polysyllabiques considérées	279
Figure 6-104	Taux de LR en corpus des adverbes monosyllabiques considérés	281
Figure 6-105	Taux de LR en corpus des formes fléchies de <i>être</i> considérées	283
Figure 6-106	Comparaison de la fréquence de LR avec le taux de réalisation de la liaison avec les formes fléchies de <i>être</i>	284
Figure 6-107	Taux de LR en corpus de la forme <i>suis</i>	284
Figure 6-108	Taux de LR en corpus des formes fléchies de <i>avoir</i> considérées	285

Chapitre 1

Introduction

La liaison est un processus, décrit comme principalement phonologique, qui met en jeu la réalisation d'une consonne orthographique, appelée consonne de liaison (désormais CL) qui, dans tout autre contexte, n'est pas phonétiquement réalisée. Dans une séquence de deux éléments consécutifs (M1 et M2), cette consonne appartient à la forme M1. Pour qu'elle soit réalisée, M2 doit commencer par une voyelle.

La liaison a fait l'objet de nombreux travaux depuis longtemps, travaux qui offrent des cadres de référence très divers (normativistes, pédagogiques, descriptifs, formalistes).

La linguistique moderne a couvert un ensemble de questions qui lui sont apparues comme étant les plus importantes et ce, aussi bien au niveau théorique qu'empirique. Ont ainsi été développés différents modèles formels, lesquels portent principalement sur des questions d'ordre phonologique et autour de l'articulation des dimensions phonologique et syntaxique. La démarche s'est voulue cumulative, elle est explicitement partie des résultats antérieurs. Au niveau méthodologique, ces approches ont récemment commencé de renouveler leurs réflexions, notamment depuis la mise à disposition de grands corpus.

L'ensemble de ces travaux a permis d'établir une description générale du phénomène que l'on peut résumer en quelques points.

Tout d'abord, on connaît les conditions potentielles de réalisation de la liaison qui sont au nombre de deux et interdépendantes (*cf.* Dell, 1973) :

(i) une contrainte d'ordre phonologique qui stipule que pour avoir un contexte de liaison il est nécessaire que la première forme (M1) se termine par une consonne non réalisée en dehors de tout autre contexte lorsque la seconde forme (M2) ne commence pas par une voyelle ;

(ii) et une contrainte d'ordre syntaxique qui permet de déterminer entre les contextes de liaison catégorique et les contextes de liaison variable : la liaison se fait catégoriquement lorsque les deux formes en présence présentent une cohésion maximale. Mais son influence semble se limiter à cela.

Ensuite, grâce au travail empirique très important, on a avancé des propositions extrêmement précises sur ce qui favorise ou handicape la réalisation de la liaison. Ces propositions font appel à plusieurs plans d'organisation et relèvent de différents domaines de la linguistique : la phonologie, la syntaxe, le lexique, la sociolinguistique, *etc.*

Ces différentes études montrent que la liaison est finalement peu réalisée et que ses contextes d'apparition se répartissent en trois classes : les liaisons catégoriques (toujours réalisées), les liaisons variables (réalisées ou non) et les liaisons erratiques (jamais réalisées). Cette tripartition fait apparaître deux points de difficulté : premièrement, on se situe dans le grammatical et / ou l'agrammatical dès que l'attention se porte sur les liaisons catégoriques et erratiques ; on se situe dans l'usage dès qu'il s'agit de considérer les liaisons variables ; et deuxièmement, le nombre de contextes considérés comme variables a considérablement augmenté à un point tel qu'il rassemble désormais la plupart des cas et que nombre de travaux se concentrent désormais sur ces derniers qui, d'un certain point de vue, manifestent la plus grande complexité.

Un autre point concerne les conditions de réalisation de la liaison. Si les dimensions ont été identifiées, le contenu des propositions relevant d'un ou plusieurs plans d'organisation peut faire l'objet de désaccords. Par exemple, est-ce la longueur de M1, celle de M2 ou encore celle de M1 + M2 qui favorise la liaison ?

Ce qui touche à la hiérarchisation de l'influence de ces dimensions demeure également complexe. Il est pourtant nécessaire de savoir de quelle façon ces dimensions

s'articulent, surtout en ce qui concerne les liaisons variables pour les raisons exposées ci-dessus. Mais, explorer chacune de ces dimensions se révèle être un exercice difficile étant donné qu'elles semblent toutes intriquées les unes par rapport aux autres. Dès qu'il s'agit d'étudier la part d'influence de l'une d'elles sur la réalisation de la liaison, on tombe systématiquement sur une autre, qui se combine éventuellement avec une troisième. Cependant, faire le point sur chacune des dimensions en jeu dans le phénomène permet de revenir sur la question des données (*cf.* ci-dessous) et de pouvoir répondre avec précision aux questions suivantes : Quelles sont véritablement ces dimensions et que recouvrent-elles ? Constituent-elles toutes réellement une contrainte ou certaines ne sont-elles pas simplement le produit du point de vue adopté ? Le fait que bon nombre d'entre elles aient été depuis si longtemps considérées comme influentes a peut-être amené les différentes études à se focaliser sur un point particulier qui, au final, ne méritait pas autant d'attention ou qui avait masqué d'autres paramètres. Comment ces dimensions s'organisent-elles les unes par rapport aux autres ? Sont-elles hiérarchisées ou simplement concurrentes selon les contextes ? Quel est le poids de chacune ?

Enfin, les données exploitées dans les différentes analyses sont souvent les mêmes, anciennes ou très spécifiques. Elles relèvent soit de travaux prescriptivistes et descriptivistes, soit de travaux portant sur une catégorie particulière de locuteurs et / ou de situations de communication (les hommes politiques avec Encrevé (1988) ; les parlars radiophoniques avec Ågren (1973) ; les classes parisiennes « cultivées » avec Delattre (1966), Fouché (1959), *etc.*).

En fait, les analyses récentes ont fourni peu d'énergie au renouvellement des données et ont majoritairement repris celles fournies par leurs prédécesseurs. Aussi, les observables sont-ils restreints à des contextes particuliers. Dès lors, toutes les généralisations antérieures doivent être réexaminées, et être confrontées à des situations socio-stylistiques plus diversifiées, plus importantes quantitativement, plus contemporaines, bref, à des données renouvelées. Ce problème avait déjà été signalé pour différentes études, notamment celles dont les auteurs n'étaient pas francophones natifs (par ex. Schane, 1968). Morin et Kaye (1982), dans la critique qu'ils font des analyses de Selkirk, notent aussi qu'une analyse globale peut être problématique selon les données que l'on utilise.

Cependant, ce type de réflexion reste marginal et la plupart des nouveaux modèles continuent de reprendre des données qui, à notre sens, ne peuvent certainement pas constituer une base empirique suffisante. Il ne s'agit pas ici d'avancer que la langue a changé à un point tel que le comportement de la liaison en a été profondément modifié. Il s'agit plutôt de considérer que, proposer une description actuelle du phénomène demande d'opérer un retour sur les données.

La possibilité de ce renouvellement passe par la mise au point d'outils permettant leur recueil et leur exploitation. C'est la démarche entreprise par le projet « Phonologie du français contemporain (PFC) » qui se propose de constituer un corpus de référence oral sur le français contemporain à même de renseigner sur la prononciation du français saisi et dans sa variation et dans la réalité de ses usages attestés.

Bien évidemment, l'entreprise est coûteuse (durée des entretiens, transcription de l'oral, *etc.*). Mais, en tout cas (et ce sont ses objectifs), elle doit permettre de (i) porter un nouveau regard sur la façon dont on travaille les données, (ii) de confronter les descriptions antérieures à ces nouvelles données. Ainsi, certains résultats qui ne trouvent aucune explication peuvent être rendus compréhensibles si l'on prend la peine de réviser ce qui « est admis par l'ensemble de la communauté » et de l'actualiser grâce aux résultats des travaux actuels. Cependant, la question de travailler avec des données uniquement attestées (ici issues de PFC) peut aussi poser problème. Comme nous le verrons notamment dans le Chapitre 6, il peut y avoir des données non présentes dans le corpus mais tout de même grammaticales, puisque le corpus a, par définition, une taille spécifiée. Cette question est d'autant plus sensible dans le projet PFC, que c'est un projet en cours, amené à évoluer. Par ailleurs, parmi les données attestées, il peut y avoir des erreurs de production, des lapsus, *etc.* Autrement dit, « attesté » ne signifie pas forcément « grammatical » (*cf.* Auroux, 1998 ; Cori et David, 2008).

Dans cette thèse, nous nous appliquerons à répondre à un certain nombre de ces questions.

Premièrement, nous tenterons de mettre à l'épreuve les différentes hypothèses et propositions que l'on a pu tenir sur les dimensions sociolinguistique (influence du registre, influence de l'âge, influence de l'origine socioculturelle, majoritairement objectivée par le niveau d'études), phonologique (qualité de la CL), syntaxique

(cohésion entre les deux termes, validité des contextes de liaison identifiés comme catégoriques et variables¹) et lexicale (influence de la longueur de M1).

Nous ferons, par ailleurs, état d'un certain nombre de questions méthodologiques : recueillir, extraire, travailler des données pour pouvoir les confronter aux hypothèses que nous souhaitons tester, demande d'être rigoureux et explicite sur différentes décisions que l'on est inévitablement amené à prendre.

Cette thèse s'organise en deux parties, dont nous présentons brièvement, ci-dessous, l'architecture.

La première partie (*Histoire des traitements de la liaison*) pose les différentes propositions qui ont été faites autour de l'analyse de la liaison :

Le Chapitre 2 (*Parcours historico-conceptuel de la liaison*) revient sur la façon dont le phénomène de la liaison a été considéré avant l'apparition des modèles formels. Ceci permettra de mieux comprendre sa complexité et pourquoi les résultats des travaux actuels ne parviennent toujours pas à proposer un traitement complet du processus.

Le Chapitre 3 (*Les dimensions d'analyse*) présente l'ensemble des dimensions impliquées dans le fonctionnement de la liaison, telles qu'elles ont été vues dans la littérature. A travers un examen critique, nous chercherons à mettre en évidence la part de l'influence de chacune des dimensions décrites. Ce sera l'occasion d'identifier précisément les propositions qui ont été avancées et qui seront confrontées à nos données dans le Chapitre 6.

Le Chapitre 4 (*Un processus phonologique*) est consacré à la dimension phonologique, dimension isolée donc par rapport à celles dont nous avons fait état dans le Chapitre 3, puisque les premiers travaux –phonologiques- formels sur la liaison (Schane, 1965) ont constitué le point de départ des différentes modélisations proposées pour la liaison. Ce chapitre présente, de façon non exhaustive, les différentes théories et modèles qui ont pu être proposés pour rendre compte du phénomène.

La seconde partie (*Analyses empiriques*) présente la méthodologie adoptée et les différentes analyses empiriques qui ont été menées :

¹ On ne traitera pas des contextes de liaisons erratiques pour des raisons exposées dans la section 6.1.

Le Chapitre 5 (*Méthodologie*) présente le cadre méthodologique dans lequel s'inscrit notre travail, c'est-à-dire le projet «Phonologie du français contemporain (PFC) : usages, variétés et structure ». Cette présentation comporte une description des objectifs de PFC et du protocole mis en place pour la collecte des données et leur exploitation. Ce chapitre décrit également le corpus que nous avons extrait et qui constitue notre base empirique.

Le Chapitre 6 (*Méthodologie, résultats et analyses*) propose une évaluation d'un certain nombre de propositions relevant des dimensions présentées dans le Chapitre 3. Une attention toute particulière sera accordée aux problèmes méthodologiques.

Première partie :

Histoire des traitements de la liaison

Chapitre 2

Parcours historico-conceptuel de la liaison

Dans ce chapitre, nous faisons état des différentes propositions qui ont cherché à caractériser le phénomène de la liaison.

Ces propositions suivent un mouvement général : on part d'un projet de simple normalisation pour aller vers une description et une compréhension du phénomène. Elles ont été suivies par des essais de modélisation, qui, le plus souvent, ont cherché à formaliser uniquement certains aspects du problème.

Les décrire, même sommairement, nous permettra en premier lieu de mieux comprendre pourquoi la circonscription du phénomène reste toujours aussi difficile.

Plus précisément, nous commencerons tout d'abord par rappeler que le problème a partie liée avec l'histoire des consonnes en ancien et moyen français (2.1). Puis, nous montrerons comment le phénomène a été successivement envisagé. Ce sont tout d'abord des enjeux normatifs qui fixent, au XVI^e siècle, son cadre d'appréhension (2.2). Cette tradition va perdurer jusqu'au XIX^e siècle et même être réactivée sous l'impulsion des pédagogues et orthoépistes au début du XX^e siècle (2.3). Mais, au milieu du XX^e siècle, un changement de cadre s'opère : des travaux délaissent la tradition normative et s'inscrivent résolument dans une perspective descriptiviste (2.4).

Nous laissons de côté, dans ce chapitre, l'exposé des propositions les plus récentes (travaux issus de la phonologie générative, travaux issus des grammaires de construction, travaux cherchant à caractériser les facteurs sociolinguistiques, *etc.*). Elles seront discutées en détail dans le chapitre « Les dimensions d'analyse » (*cf.* Chapitre 3).

2.1 Rappel historique

Si le phénomène de la liaison est difficile à comprendre, c'est à cause de l'histoire des consonnes. On sait (Bourciez, 1945 ; Fouché, 1952 ; Brunot, 1966 ; Bonnard, 1979) que les consonnes de liaison correspondent à des consonnes « historiques » (cf. Laks, 2005b : 112) qui étaient fixes à une époque où, conséquemment, la liaison n'existait pas. Nyrop (1934) écrira que la liaison est :

[...] souvent un phénomène de réaction : c'est la conservation ou la résurrection, dans un cas spécial, d'un phonème qui est devenu muet partout ailleurs, d'une lettre morte. (Nyrop, 1934 : 127)

Un peu plus tard, Grevisse (1969 : 47) expliquera lui aussi que la liaison résulte « [...] d'un usage ancien, selon lequel les consonnes finales se prononçaient toutes », tout en précisant qu'elle :

[...] n'est donc que la persistance d'un ancien enchaînement. Des adjectifs comme *grand* et *long* s'écrivaient *grant* et *lonc* conformément à leur prononciation, laquelle s'est donc maintenue dans la liaison.

Selon ces auteurs, la liaison correspond à une sorte de résurrection phonétique des consonnes. Pour expliciter cette « résurrection », il est nécessaire d'opérer un retour en arrière de quelque dix siècles, à l'époque de l'ancien français (IXe – XIIIe siècles), période pendant laquelle toutes les consonnes écrites se prononcent, la graphie ayant une forme beaucoup plus proche de la phonie que l'orthographe actuelle. Le *-t* final, par exemple, est prononcé [θ] dans des mots comme *aimet*. Tous les *-s* du pluriel se font entendre (par ex. *les omes* /les hommes/ prononcé [lezomœs]), de même que toutes les finales en *-z* qui ont valeur d'affriquée (par ex. le *-z* dans *chantez* a la valeur [ts]).

C'est entre le IXe et le XIVe siècle que l'on assiste à une évolution de la prononciation, évolution qui se traduit par l'amuïssement progressif mais certain des consonnes finales. Cet affaiblissement touche en premier lieu les consonnes dont le caractère phonétique est faible, comme les dentales. C'est principalement au cours du

XIII^e siècle que le système de l'ancien français se simplifie et que l'amuïssement se généralise à toutes les consonnes finales (par ex. *grand* prononcé *grant* [grāt] en ancien français devient *gran* [grā]²). Au XIV^e, seules les consonnes finales à la pause sont conservées. Enfin, au XVI^e siècle, la prononciation des consonnes finales a quasiment disparu, excepté lorsqu'elles sont suivies d'une voyelle ou qu'elles se situent en fin de phrase.

Cette disparition trouve sa motivation dans une tendance très forte à la syllabation ouverte (Martinet, 1964 : 328 ; Pulgram, 1970 : 69), dont Laks (2005b : 122) dira qu'elle « *constitue un fait de structure massif qui perdure depuis des siècles* ». Citant Väänänen (1962 : 129), il fait remarquer que :

[...] la chute des consonnes finales correspond à une dynamique très profonde qui trouve son origine dès le latin archaïque. L'inversion du rapport ancien qui voit l'opposition de hauteur vocalique supplanter l'opposition de longueur y libère l'accent quantitatif, affaiblissant les syllabes post-toniques. En latin tardif, la tendance à l'affaiblissement des consonnes finales se renforce encore et se généralise. (Laks, 2005b : 112)

Delattre (1966) avait lui aussi rappelé que le haut degré de syllabation ouverte, qui caractérisait la prononciation de la fin du Moyen Age, était la cause de la conservation des consonnes finales dans la liaison entre deux mots étroitement unis :

[...] dans la chaîne verbale, la consonne finale du premier mot était tellement unie (syllabiquement) à la voyelle initiale du mot suivant et tellement séparée (syllabiquement) de la voyelle précédente, qu'elle n'a pas été atteinte par l'amuïssement de la même consonne finale dans le mot isolé ou suivi de consonne [...]. (Delattre, 1966 : 55)

Ce changement, qui se fera sur le long terme et qui prendra fin au XVII^e siècle, ne sera jamais totalement accompli. Comme le rappelle Laks (2005b : 113), les marques

² En accord avec la règle de dévoisement des occlusives finales datant du VIII^e siècle, *grand* était prononcé [grāt], conformément à sa graphie. Si la consonne fixe disparaît au niveau phonique, elle reste graphiquement présente jusqu'à la fin du XII^e siècle. L'évolution de l'orthographe aidant, le mot s'écrira par la suite avec un <d> final, conformément d'une part à son origine latine *grandis* et d'autre part à ses nombreux dérivés lexicaux (*grande*, *grandir*, *grandement*, etc.).

morphologiques du pluriel, tel que le *s* par exemple, ont particulièrement résisté à ce traitement³.

Parallèlement à cette évolution phonétique qui touche en premier lieu les couches populaires pour s'étendre peu à peu à la langue des plus érudits et qui continue d'évoluer, l'orthographe commence à se fixer (XVe siècle). Les grammairiens (par ex. Meigret, 1542) tentent alors de restituer ces consonnes disparues. Ce travail se fera pas à pas, d'une part, par le biais des poètes et des règles de la versification et, d'autre part, par les tentatives de stabilisation et de normalisation de la langue française. La réintroduction des consonnes finales orthographiques (XVIIe siècle), réalisée de manière plus ou moins anarchique⁴, ne fera qu'accentuer le clivage entre les plus érudits (qui pratiquent le bon usage) et les « masses populaires ». Comme aiment le rappeler les spécialistes de la pédagogie contemporaine, la graphie ne rend plus compte des usages et l'orthographe n'est plus la langue, elle n'en est plus que « l'habit ».

Examinons maintenant les différentes étapes qui ont marqué l'analyse de la liaison.

2.2 Les normativistes

Brunot (1966) retrace entre autres l'évolution du phénomène de la liaison et fait remarquer que les règles prescriptives ayant trait à la réalisation des consonnes finales en liaison ont été l'objet de nombreuses réflexions et ce, depuis le XVIe siècle. En effet,

³ Laks (2005b : 113) propose, à côté de cette première dynamique constituée par la tendance à la syllabation ouverte, une seconde dynamique, celle du marquage morphologique du nombre, contradictoire à la première. Il fait remarquer que « *cette dynamique syllabique est, dès l'origine, étroitement corrélée à des phénomènes rythmiques et accentuels et se marque par une désautonomisation du mot au profit de la construction de groupes phonétiques liés par l'enchaînement,* » mais qu'elle n'atteindra jamais « *son terminus ad quem* ». Si la plupart des consonnes finales disparaissent, il en est une qui résiste particulièrement à l'amuïssement : le *s* à valeur structurale, « *préfigurant ainsi le fonctionnement morphologique du marquage du nombre préposé* » (observé par Dauzat (1930 : 97), Bourciez et Bourciez (1967 : 123), Väänänen (1962 : 128)).

⁴ La première tendance, motivée par les impératifs de l'écriture manuelle (par ex. l'écriture judiciaire de l'administration royale), a consisté à rendre les mots plus lisibles (particulièrement les monosyllabes devenus très nombreux). La seconde tendance, due au besoin de clarté des imprimeurs, a préféré un système de voyelles accentuées plutôt qu'une suite sans fin de consonnes.

si l'usage était fixé pour un certain nombre de cas, d'autres liaisons restaient irrégulières (par ex. la restitution du [z] final des noms ou des adjectifs qui ne s'entendaient plus à la pause). Et d'autres encore relevaient, selon les grammairiens de l'époque (par ex. Thurot, 1881, (2) : 61), de l'anarchie la plus complète (par ex. *on-z-a*, *un laid-z-homme*, Cauchie, 1586 ; *j'ai-z-eu*, *il a-t-ouïe*, *je n'en ai point-n-eu*, Hindret, 1687 ; *avant-z-hier*, Vaugelas, 1647)⁵.

Brunot (1966) souligne que les locuteurs ont depuis tout temps évité les hiatus et que l'évolution de la langue a provoqué un déclin important de la prononciation des consonnes finales :

[...] il semblerait que le peuple avait toujours une répugnance naturelle aux hiatus et que, pour les résoudre, il intercalait entre les deux voyelles une consonne, sans se soucier de l'ancienne consonne finale. [...] il s'est fait un compromis entre l'ancien usage où la consonne finale s'articulait avec la voyelle initiale de n'importe quel mot suivant et les tendances populaires récentes qui allaient à l'amuïssement général de la consonne. La date où ce fait s'est produit peut être fixée au second quart du XVII^e siècle.

Les réflexions de Brunot amènent facilement à conclure que la liaison a toujours fait partie des usages, même si son fonctionnement semble anarchique. Le « compromis » qu'il propose, fait état d'un équilibre trouvé entre deux tendances contraires : d'une part la disparition de la plupart des consonnes finales et d'autre part, la prononciation systématique de ces mêmes consonnes devant tout mot à initiale vocalique. Cette régularisation a trouvé sa source dans le système normatif mis en place au cours du XVII^e.

Ce système, dont nous allons décrire les grands principes, s'est fixé pour objectif de structurer et de stabiliser la langue, autrement dit, de la fixer. Des règles précises d'usage, à savoir les obligations et les interdits de la langue, ont ainsi été établies et sont devenues, par la suite, plus ou moins communes à l'ensemble des locuteurs natifs grâce, principalement, à l'enseignement.

⁵ Pour une description complète, voir Brunot (1966 (IV) :1).

Au XVII^e siècle, la liaison n'est certainement pas considérée ni même envisagée comme un problème de langue en soi, mais bien plus comme un phénomène délimitant les classes sociales. Comprendre les mécanismes qui la régissent ne fait pas partie des ambitions des normativistes, qui sont les premiers à avoir posé les principes de fonctionnement de la liaison et à les avoir diffusés dans la communauté. La tâche principale des intellectuels de l'époque consiste donc non pas à expliciter les causes de la réalisation de la liaison mais à indiquer dans quelles circonstances il faut ou non la réaliser.

Cette période marque un tournant décisif pour le français, qui a quasi définitivement pris le pas sur le latin. Les enjeux intellectuels de l'époque visent alors à rendre la langue pure et éloquente, tout en lui fournissant une grammaire de référence et un statut de langue véhiculaire⁶. Les règles normatives adoptées s'inscrivent entièrement dans des considérations subjectives, esthétiques et sociales, et sont tout à la fois arbitraires et stables.

Il est évident que la liaison n'a pas été épargnée par la mise en place de ce système rigoureux, qui s'est évertué à proscrire tout ce qui n'était pas conforme à cette norme « enfin » établie.

Avec la publication de son ouvrage *Les Remarques sur la langue française*, Vaugelas (1647) constitue la codification d'un usage linguistique dont l'objectif est de devenir le modèle du français écrit (en prose et en poésie) et de régler les habitudes linguistiques. La langue étant soumise à des changements constants, Vaugelas détermine un usage ayant pour fondement la façon de parler d'un groupe de locuteurs à registre unique, « composé de l'élite des voix », à savoir la Cour (1647, Préface a, II : 2).

Le modèle nouvellement créé fixe certains usages « actuels » et délimite ce qu'il « ne faut pas se laisser prononcer » des impératifs relevant d'un contentement général. C'est à partir d'une norme implicite, socialement déterminée par un groupe de locuteurs, qui n'est pas majoritaire, que Vaugelas pose une norme tout à fait explicite. Elle continuera à s'imposer comme un modèle idéal à suivre. Cette codification du bon usage survivra

⁶ Sur les enjeux intellectuels de l'époque, voir D. Trudeau, 1992, *Les inventeurs du bon usage, 1529-1647*. Paris : Editions de Minuit.

et se renforcera au travers de la pratique pédagogique française. Le caractère évolutif de la langue parlée, si cher à Vaugelas, ne sera pas retenu par ses successeurs, et l'usage qui a été décrit à cette époque sera fixé par une image écrite qui deviendra une norme à adopter. Ce système de codification reposant sur un registre univoque et donc non accessible à toutes les classes sociales, sera érigé au rang de registre exemplaire à suivre par tous.

Fixer le « bon usage » est longtemps resté une des motivations principales des grammairiens. L'orthoépie, le choix du vocabulaire, la morphologie ou la syntaxe ont servi à définir une norme retenant seulement certaines façons de parler et rejetant les autres, considérées alors comme incorrectes.

2.3 Les prescriptivistes

La norme établie par les normativistes verra sa pérennité assurée par les pédagogues orthoépistes et phonéticiens de la première partie du XXe siècle.

Les prescriptivistes s'adressent, pour la plupart, à un public natif francophone averti ou apprenant avancé. On peut par exemple citer Martinon (1913), lequel pose très clairement ses objectifs dans l'introduction de son traité de phonétique intitulé *Comment on prononce le français* (Préface, II). Dans cet ouvrage, Martinon s'adresse de préférence à un public de francophones natifs ou du moins très au fait de la langue. De la même façon, dans le *Bon Usage*⁷, Grevisse (1969), grammairien normativiste, vise un public désireux de parfaire sa maîtrise de la langue.

Le travail des pédagogues consiste en grande partie à fournir un modèle de « prononciations agréables à l'oreille » (Martinon, 1913 : I) adapté à l'époque qui est la leur, où le locuteur ne « commettra pas de faute » (Fouché, 1959 : 3). Dès lors, légitimer

⁷ *Le Bon Usage*, qui constitue l'ouvrage principal de Grevisse (1969), ne compte pas moins de treize éditions et fait encore, aujourd'hui, référence.

l'usage des tournures se fera pour les uns par le biais des « *bons écrivains* » ceux qui ont « *le souci de bien s'exprimer* » (Grevisse, 1969 : 6), remplaçant ainsi « *la partie la plus saine de la Cour* », et pour les autres, par les locuteurs issus des classes cultivées de Paris (entre autres Martinon, 1913 ; Nyrop, 1934).

Si certains de ces prescriptivistes, comme par exemple Grevisse (1969), reconnaissent que la norme, trop souvent théorique, peut être arbitraire, ils estiment qu'il faut tout de même s'élever contre le laisser-aller. Ces auteurs perpétuent ainsi la tradition normative en promulguant ce qui doit être et ce qui ne doit pas être.

Le schéma explicatif du phénomène ne fait pas partie de leurs desseins. On peut citer les travaux de Fouché (1959), qui décrit de façon très précise des niveaux stylistiques distincts (la conversation courante et le style soutenu), ou encore Grevisse (1969) qui utilise une terminologie tripartite (liaisons obligatoires, liaisons recommandées et liaisons inusitées).

D'après Durand et Lyche (2003 : 260) le problème majeur que posent les travaux de ces linguistes est la quasi impossibilité « [de] *séparer le prescriptif du descriptif* » tant l'influence de la norme sociale de l'époque est prégnante :

Le plus souvent, les milieux sociaux et les styles décrits ne sont ni identifiés de façon précise, ni stables. En décrivant la liaison, des phonéticiens dont on admire l'objectivité dans d'autres domaines, sont pris dans la tenaille de la norme et ne peuvent cacher leur mépris ou leur irritation à l'usage de diverses classes sociales. Passy (1913) et Martinon (1913) ne sont ni les premiers ni les derniers [...]. (Durand et Lyche, 2003 : 260)

La méthode utilisée par les prescriptivistes pour étudier la liaison reste strictement identique. Il faut souligner que la plupart sont phonéticiens (Martinon, 1913 ; Grammont, 1914 ; Fouché, 1959) ; ils décrivent le phénomène d'un point de vue phonétique. Martinon (1913), par exemple, expose les conditions de réalisation de la liaison en décrivant les consonnes de liaison selon le groupe phonétique auquel celles-ci appartiennent. Ainsi, il traitera les labiales (*p*) et les gutturales (*c*, *g*) dans un groupe et les dentales (*d* et *t*) dans un autre. Chacune des catégories proposées sont accompagnées de commentaires ayant trait aussi bien à l'histoire de la langue, qu'à la stylistique : « *c'est ici qu'intervient le goût personnel pour les facultatives* » (Martinon, 1913 : 352).

Le comportement de liaison est donc décrit sous forme de règles, pour corriger la prononciation ou au mieux, la perfectionner. Ainsi, si les descriptions fournies par les descriptivistes sont principalement d'ordre phonétique, elles n'en demeurent pas moins normatives. Les données empiriques utilisées résultent invariablement des locuteurs jugés les plus « cultivés », le plus souvent des Parisiens. Pour Grammont (1914), le critère de référence utilisé sera « [...] *la prononciation de la bonne société parisienne, constituée essentiellement par les représentants des vieilles familles et de la bourgeoisie* » (Grammont, 1914 : 1). Fouché (1959 : II-III), quant à lui, se réfère aux données résultant d'une enquête longitudinale sur « *la prononciation soignée en usage dans une conversation soignée chez les Parisiens cultivés [...] nés vers la fin du XIXe siècle ou plus tard.* ». Encrevé (1988) rappelle d'ailleurs que :

Ces travaux ne proposent aucune théorisation ni formalisation mais font « autorité » : [...] parce qu'ils sont en continuité avec les travaux des phonéticiens et orthoépistes antérieurs. (Encrevé, 1988 : 87)

Les observations de Martinon (1913) lui permettront de mettre en valeur un certain nombre de principes (stylistiques, sémantiques) concernant les liaisons facultatives. Lorsqu'il s'agit d'indiquer quand faire une liaison, exclusion faite de celles qui sont « définitivement obligatoires », Martinon propose d'une part, que les liaisons ne peuvent opérer qu'entre deux mots *liés par le sens* à l'intérieur d'une même phrase, d'autre part, que même si les liaisons facultatives sont très sûrement livrées au goût de chacun, elles restent dépendantes de l'intention communicationnelle :

[...] il est évident qu'on en fait plus en lisant qu'en parlant, parce qu'en lisant on recherche la correction du langage, [...] tandis qu'en parlant on ne cherche qu'à se faire comprendre avec le moins d'effort possible ; on en fait plus aussi dans un discours suivi, pour le même motif, que dans une conversation familière. (Martinon, 1913 : 356)

Un peu plus tard, Fouché (1959 : 434-477) remarque que la liaison est le seul lieu où il est nécessaire d'opposer la conversation soignée et la conversation familière. Pour chacun de ces styles, il propose ce qui peut être admis ou permis ainsi que ce qui est proscrit.

Et comme le souligne Encrevé (1988 : 44), « *Fouché ne pratiquait pas l'analyse quantitative d'un corpus fermé, [...]* » mais « *construit une moyenne qualitative, c'est-*

à-dire normative, prescriptive [...] » lui permettant, à la suite de Vaugelas (1647), d'indiquer un bon usage :

Loin de nous la pensée que telle ou telle prononciation passée sous silence ne soit pas bonne. Mais on ne commettra pas de faute en s'en tenant à celles qui sont notées ici. (Fouché, 1959 : 3)

Comme le rappelle Encrevé (1988 : 87), l'apport majeur des prescriptivistes consiste en un considérable travail de description phonotactique, lequel a abouti à un classement rigoureux des contextes de liaison et à une étude détaillée de la liaison selon « *le niveau de langue employé* ».

Parallèlement pour une part à ces travaux, où l'importance de la norme reste prépondérante, on assiste au développement d'une nouvelle démarche, qui consiste à privilégier la description plutôt que la prescription.

2.4 Les descriptivistes

Nyrop (1934) introduit son *Manuel de phonétique du français parlé* par la question suivante :

Pourquoi aussi vouloir établir un seul bon usage, quand la pratique nous apprend tous les jours qu'il y en a plusieurs ? (Nyrop, 1934 : VI)

Cette question illustre bien le nouveau mode d'appréhension du phénomène, que l'on peut associer plus ou moins à la seconde moitié du XXe siècle. Cette période est en effet marquée par une prise de conscience – ou plutôt une reconnaissance de l'existence – de la diversité des usages linguistiques et de l'existence d'une mécanique linguistique complexe dans le phénomène de la liaison. Mais elle est également marquée par le rejet de l'attitude normative de la grammaire traditionnelle, et s'assigne comme objectif premier de comprendre les phénomènes langagiers, de rendre compte de la structure de la langue, plutôt que de prescrire des règles. Si la plupart des travaux (Nyrop, 1934 ; Delattre, 1966 ; Léon, 1966) continuent d'être influencés par la norme, ils tentent de s'en déprendre ; et leurs descriptions, riches et pertinentes, continuent d'être utilisées par la linguistique moderne. Certains auteurs y parviendront (par ex. Ågren, 1973 ;

Lucci, 1983) grâce à leurs analyses empiriques sur les liaisons facultatives et fourniront à la linguistique contemporaine une source d'inspiration très importante⁸.

Pierre Delattre est l'un des premiers à comprendre que les mécanismes linguistiques sont plus puissants qu'une norme sociale et pédagogique et qu'ils constituent le pivot du fonctionnement de la liaison. Il rappelle que le phénomène de liaison est une survivance d'enchaînements de consonnes finales de l'ancien français. Ces consonnes désormais purement graphiques ne trouvent de réalisation que dans certaines situations particulières de la chaîne parlée : lorsque l'union entre les deux termes, garantie par l'usage, est assez forte.

Dès 1947, dans son article intitulé « La liaison en français, tendances et classification », Delattre établit une grille descriptive de tous les contextes de liaison potentiels qu'il répartit en trois catégories de contextes : ceux où la liaison est obligatoire, ceux où elle est facultative et enfin ceux où elle est interdite.

Suite à ce travail dans lequel il fait aussi état de différents styles possibles, Delattre (1955) publie un article uniquement consacré aux liaisons facultatives : « La fréquence de la liaison facultative en français ». Il y insiste sur le fait que les distinctions établies ne sont pas absolues et qu'il peut y avoir de la variation selon le style employé. L'examen des liaisons facultatives l'amène en outre à constater que l'on peut aussi bien les faire que les omettre sans pour autant dépasser les bornes du langage accepté. La fréquence avec laquelle ce type de liaison est réalisé est le résultat d'un conditionnement multifactoriel. Les quatre facteurs les plus importants sont respectivement et par ordre d'importance : la stylistique, la syntaxe, la prosodie et la phonétique.

Delattre mettra ainsi en évidence :

- (i) l'influence de la situation de communication et du registre employé (stylistique) ;
- (ii) l'influence de la syntaxe, les liaisons facultatives étant en partie réglées par le degré de solidarité des mots en contact ;

⁸ On peut notamment citer Bybee (2001a et b, 2005), laquelle s'inspirant des travaux de Ågren (1973) développera un modèle cognitiviste de la liaison, fondé sur l'influence de la fréquence lexicale (cf. 4.4).

- (iii) l'influence de la prosodie, à savoir la longueur des éléments à lier, l'intonation de la phrase, l'accent d'insistance, etc. ;
- (iv) et l'influence du degré phonématique.

Delattre n'est pas le premier à avoir donné une description multidimensionnelle du phénomène. Il est cependant le premier à regrouper tous ces facteurs dans une démarche qui se veut explicative et à tenter de mettre au jour une hiérarchie dans la fréquence des liaisons facultatives. Tous ses articles (1947, 1955, 1966)⁹ portant sur la question, et plus particulièrement son étude sur les liaisons facultatives, visent à en fournir une description exhaustive.

A l'instar de ses prédécesseurs, et dans une perspective pédagogique¹⁰, il reconnaît à la liaison son caractère social et normé, qu'il inscrit parmi les facteurs stylistiques. La stylistique constitue, selon Delattre, le critère le plus déterminant dans la réalisation de la liaison facultative.

Comme la catégorie des facultatives regroupe des contextes très différents, la liaison est très fréquente dans certains cas (*il commençait_à lire*) et très rare dans d'autres (*les pauvres_attendront*). Delattre (1955) explique que le choix du style, exigé par les circonstances, peut faire varier considérablement la fréquence des liaisons. Pour synthétiser, plus le registre de langue est élevé, plus les liaisons facultatives ont tendance à être réalisées. Et inversement, les liaisons sont d'autant moins réalisées que le style est familier.

Au-delà de son rôle de pédagogue, Delattre essaie de comprendre et de rendre compte de la totalité des processus linguistiques en jeu dans la réalisation de la liaison et reconnaît ainsi implicitement l'existence de la variation. Dans la mise en avant de l'idée de variation, totalement ignoré par les recherches qui suivront (*i.e.* les générativistes : Schane, 1967a et b ; Dell, 1970 et 1973 ; Selkirk, 1972 ; *etc.*) et qui lui préféreront les travaux normativistes (Fouché, 1959 ; Grevisse, 1969 ; Grammont, 1933), Delattre est un précurseur.

⁹ L'ensemble des recherches de Delattre sur la liaison se trouvent réunis dans son ouvrage intitulé *Studies in French and Comparative Phonetics* (1966).

¹⁰ Rappelons que si Pierre Delattre est reconnu en tant que phonéticien, ses travaux sont destinés en priorité à un public d'apprenants du français langue seconde et à des étudiants d'orthophonie (*cf.* notamment Delattre, 1951).

Dans la lignée de Delattre, on peut citer l'étude d'Ågren (1973) lequel, comme nous l'avons souligné précédemment, s'est lui aussi efforcé de comprendre le fonctionnement des liaisons. Critiquant ses prédécesseurs sur leur objectif purement normatif et reconnaissant à Delattre le statut de pionnier en matière de tentative de compréhension du phénomène, il dira dans l'introduction de son ouvrage *Etude sur quelques liaisons facultatives dans le français de conversation radiophonique* :

J'ai pu constater que tous les manuels de phonétique, Martinon, Grammont, Fouché, Bruneau, Nyrop, Peyrollaz, et j'en passe, se sont occupés de la liaison dans le seul but pratique et normatif¹¹. En se limitant à cette seule perspective, ils ont plus ou moins négligé ou fait passer à l'arrière-plan les facteurs sous-jacents de la liaison et de la non-liaison. C'est Delattre, le premier dans les années cinquante, qui s'efforce de rassembler les informations éparses sur le pourquoi de la liaison facultative. (Ågren, 1973 : 14)

A l'instar de Delattre, Ågren considère que la liaison facultative est le produit de la combinaison de multiples facteurs. Beaucoup plus dégagé de la norme que son prédécesseur, Ågren rend compte, dans la perspective qui est la sienne, de l'existence de nouveaux facteurs (par ex. la fréquence d'emploi) qui doivent être considérés. Cependant, il fait état de la complexité de cette entreprise. Ces facteurs restent difficilement cernables du fait même de leur interdépendance :

A part les facteurs « contrôlables » qui viennent d'être décrits, il en existe quelques autres, dont l'influence est moins facile à vérifier. [...] Cependant, lorsque l'on entame leur description, il faut être pleinement conscient que l'on quitte le domaine de l'objectif pour entrer dans celui du probable. C'est dire que presque rien de ce qui sera avancé ne pourra être confirmé d'une façon entièrement satisfaisante. (Ågren, 1973 : 28)

¹¹ Ågren (1973) appuie ses propos sur un article de Mok (1966 : 27-39) qui reproche aux différents spécialistes de phonétique ou de prononciation française de ne toujours traiter que les mêmes aspects de la liaison, à savoir les différents niveaux d'usage de la langue. Les auteurs sont unanimes quant au fait que la fréquence de la liaison varie selon les différents niveaux d'usage de la langue. Par exemple, pour Fouché (1959), c'est cette fréquence même qui a été le critère principal sur lequel il a fondé sa distinction de quatre sortes de français : débit de conversation courante, soignée, conférence et diction. En outre, Ågren critique ouvertement la méthode de travail et de recueil de données selon laquelle le phénomène est étudié et qui reste fondamentalement identique : « *Descriptive en même temps que normative et presque exclusivement centrée sur le locuteur. Le comportement, en matière de liaison, du locuteur, qui sera le plus souvent le phonéticien lui-même, est décrit en forme de règles, destinées à des locuteurs du français, en vue d'une correction de prononciation, à des étrangers qui veulent apprendre à bien parler cette langue.* ».

L'objectif visé par Ågren consiste à tester chacun des facteurs mis en évidence par le biais d'un corpus strict. Il différenciera ces facteurs en deux groupes : le premier fera écho aux facteurs internes (syntaxe, phonétique, prosodie, etc.), qu'il dira « contrôlables » ; le second aux facteurs externes (stylistique, socio-démographiques, etc.) qui comportera l'ensemble de ceux qu'il tient pour subjectifs (par ex. la fréquence d'emploi).

2.5 Bilan du parcours historico-conceptuel

Il faudra attendre la seconde moitié du XXe siècle pour que les travaux sur la liaison se déprennent des enjeux normatifs qui ont été constitutifs de son cadre originel d'appréhension. Cependant, ce qui reste remarquable dans ces travaux antérieurs, et plus particulièrement ceux de la fin du XIXe et du début du XXe, c'est la double lecture qu'ils continuent d'autoriser. En d'autres termes, le « discours normatif » n'a pas complètement hypothéqué la lecture des données empiriques sur lesquelles ils se sont appuyés. Et il faut souligner le fourmillement de notations, d'indications, voire de descriptions qu'ils ont fournies. Certes, on ne sait plus toujours distinguer ce qui relève du « prescriptif » de ce qui relève du « descriptif », mais c'est sur ce fond empirique que les travaux descriptivistes ont pu voir le jour.

Avec les descriptivistes, et notamment avec Delattre, c'est la multidimensionnalité du phénomène qui est mise de l'avant. Sur un fond empirique tout aussi important, leurs travaux ont permis de mettre au jour l'essentiel des dimensions d'analyse, qui seront retravaillées par les linguistes contemporains :

- caractérisation de la consonne de liaison ;
- importance du paramètre de la longueur ;
- inventaire des contextes syntaxiques ;
- caractérisation des niveaux de langue ;

Cette multidimensionnalité est appréhendée d'emblée comme une complexité, car on reconnaît que les dimensions sont interdépendantes.

Le second apport majeur des descriptivistes concerne la notion de variation. C'est sur ce point qu'ils se démarquent radicalement de leurs prédécesseurs prescriptivistes. Il n'y a plus une seule et bonne manière de parler, il peut y en avoir plusieurs. Aussi la variation observée dans la réalisation de telle ou telle liaison n'est plus « une faute » mais fait partie intégrante du phénomène dont il faut rendre compte. Et cette variation contribue elle aussi à la complexité du phénomène, du fait du nombre de paramètres en jeu et de leur interdépendance.

Cependant la variation à laquelle ces travaux s'attachent est plutôt orientée vers la variation stylistique et les nécessités de la langue. La part de spontané du locuteur n'est quasiment pas représentée, autrement dit la part de la variation inhérente n'est pas reconnue. Et les résultats des descriptivistes restent assez vagues : on conclut souvent en reconnaissant la difficulté et la complexité à appréhender le phénomène de la variation. Malgré tout, il importe de reconnaître que ces travaux sont à l'origine d'une nouvelle problématisation de l'étude de la liaison, soit l'étude des facteurs influençant la réalisation de la liaison. Nous insistons sur ce point, car, nombre de travaux ultérieurs, comme nous le verrons dans le chapitre suivant, se sont d'abord attachés à décrire l'une ou l'autre de ces dimensions, rarement plusieurs, et jamais à rendre compte du poids ou de l'influence de chacune d'entre elles. C'est en ce sens que la perspective des descriptivistes demeure résolument « moderne », même si les formulations de l'époque restent encore dans l'implicite.

Enfin, il faut souligner l'apport des prescriptivistes et des descriptivistes sur le plan empirique. Certes, le reproche qu'on peut leur adresser est d'avoir accumulé, classé, analysé des données sans que l'on puisse en déduire un quelconque système linguistique général¹². Ce sont d'abord et avant tout leurs données qui ont été (re)testées, retravaillées. Et qui continuent de l'être¹³ (cf. par ex. Bybee, 2001 : 167-188).

¹² Voir Mounin (1968). *Clefs pour la linguistique*. Paris : Seghers.

¹³ La mise au point de corpus et d'outils informatiques, depuis une vingtaine d'années, depuis près de dix ans avec le projet PFC, est en train de changer la donne. C'est en tout cas l'objectif.

Chapitre 3

Les dimensions d'analyse

Ce chapitre a pour objectif de décrire les processus en jeu dans le fonctionnement de la liaison tels qu'ils sont traditionnellement¹⁴ dépeints depuis les travaux des descriptivistes. Nous tenterons ainsi de dégager les principes définitoires les plus communément admis, et de rendre explicite les différentes facettes du phénomène.

La réalisation de la liaison en français apparaît, comme nous l'avons fait remarquer dans le chapitre précédent (*cf.* 2.4), comme étant contrainte par de nombreuses dimensions qui relèvent de plans d'organisation différents, ce qui rend son analyse complexe. Toute la littérature fourmille de descriptions éparses. Ont été principalement notées et étudiées, les dimensions phonologique, morphosyntaxique et lexicale. Plus récemment, grâce aux travaux d'Encrevé (1988), le caractère éminemment variationnel de la liaison a été mis en avant et la sociolinguistique s'est attachée à essayer d'en comprendre la dynamique.

La présentation qui va suivre, s'organise autour de ces différents points. Dans un premier temps (*cf.* 3.1), nous nous attacherons aux aspects phonologiques (enchaînement, (re)syllabation, nature des consonnes de liaison, hiatus). Dans un second temps, l'analyse de la dimension morpho-syntaxique (*cf.* 3.2) nous amènera à considérer la tripartition des contextes de liaison mise en place et à discuter de son contenu. Nous verrons, dans un troisième temps (*cf.* 3.3), quel a été l'apport de la dimension sociolinguistique dans la caractérisation de la variation. Enfin, dans un quatrième temps

¹⁴ Par *traditionnellement*, nous entendons tout ce qui fait l'objet d'un consensus large, depuis les descriptivistes jusqu'à la période post-généraliste. Il n'en demeure pas moins que sur un certain nombre de points (par ex. le hiatus), les analyses sont divergentes. Dans ces cas, nous ferons état des discussions que ces différents points ont suscité.

(cf. 3.4), nous aborderons l'ensemble des facteurs lexicaux également susceptibles de contraindre la réalisation de la liaison (longueur et fréquence d'apparition des éléments à lier).

3.1 Phonologique

3.1.1 Principes généraux du phénomène

Le français lie, unit, enchaîne les mots les uns à la suite des autres, resyllabant, usant de stratégies de « réparation » pour les mots commençant par une voyelle (par ex. la liaison, l'épenthèse). La liaison est un phénomène de sandhi externe ; elle fait apparaître phonétiquement une consonne finale qui, généralement, appartient orthographiquement au mot, mais qui n'est jamais réalisée en-dehors de ce contexte.

Selon Fouché (1959 : 434), la liaison relève de la phonétique syntactique et « consiste par conséquent à prononcer un mot commençant par une voyelle une consonne finale, muette en dehors de cette condition ». Plus formellement, la liaison peut être vue comme un processus phonologique caractéristique du français dans lequel la consonne finale de certains mots (par ex. le *t* final de *petit*), consonne de liaison (CL¹⁵), apparaît en surface quand :

- (i) le mot suivant commence par une voyelle ;
- (ii) et lorsque les deux mots sont étroitement liés syntaxiquement (par ex. *un petit ami* [ɛ̃pətitamɪ] / *un petit apprend vite* *[ɛ̃pətitaɾãvit(ə)]).

Dans la plupart des cas répertoriés, les consonnes de liaison sont enchaînées, c'est-à-dire réalisées comme attaque de syllabe (par ex. *vos amis* [vo.za.mi]). Pour une description phonétique traditionnelle, il s'agit d'un fait de « resyllabation » qui

¹⁵ Désormais, lorsque nous utiliserons l'abréviation « CL », il faudra entendre Consonne de Liaison.

intervient lorsque « la consonne finale du premier mot se soude à la voyelle initiale du second pour former syllabe avec elle » (Fouché 1959 : 434). Ainsi, lors du processus de liaison, la « resyllabation [...] fait entendre une CL à l'attaque de la première syllabe du second mot en jeu » (Encrevé, 1988 : 23). Ce phénomène est désigné sous le terme d'enchaînement.

Toutefois, en opposition à la resyllabation, il est nécessaire d'admettre l'existence d'un non-enchaînement qui constitue le propos central de l'ouvrage d'Encrevé (1988). En effet, bien que l'enchaînement soit représentatif de la plupart des cas de liaisons, il arrive aussi que la consonne en question soit réalisée en position de coda (*vos amis* [voʒ.a.mi]). On parle alors de liaison non enchaînée :

[...] enchaînement et non-enchaînement peuvent se constater en français dans la chaîne parlée pour toutes les consonnes placées devant un mot à initiale vocalique, que ces consonnes soient toujours prononcées (Consonnes Fixes, dorénavant notées CF) ou qu'elles ne s'entendent que dans ce contexte particulier (CL). (Encrevé, 1988 : 29)

Ainsi, pour réaliser une liaison, il faut une consonne finale inexistante phonétiquement dans tout autre contexte que celui de la liaison.

3.1.2 Une consonne de liaison

En français, on peut distinguer deux grands types de consonnes finales : les consonnes dites *fixes* et les consonnes de *liaison*. Les consonnes finales fixes¹⁶ sont par définition toujours prononcées (par ex. *donc* [dɔ̃k], *grandir* [grɑ̃dir], *mal* [mal]). Elles constituent donc un phonème comme un autre dans le mot.

¹⁶ Tranel (2000 : 48) préférera l'utilisation du terme *intrinsèque* à celui de *fixe*, l'étude de Morin (1998) ayant montré que ces consonnes pouvaient parfois *manquer de fixité phonétique*. On peut en effet observer une certaine fragilité des liquides en finales lorsqu'elles sont précédées d'une obstruante (les groupes *obstruantes* + *liquides* désormais notés OL).

La réalisation des consonnes de liaison subit, quant à elle, des contraintes phonologiques, syntaxiques et stylistiques¹⁷ (par ex. *petit* [pəti] + *enfant* [ãfã] sera, prononcé [pətiãfã]).

3.1.2.1 Bref inventaire

L'inventaire des consonnes de liaison est rapide. On en dénombre sept au total : [n], [g], [k], [p], [t], [ʀ] et [z], les plus productives étant [z], [t] et [n]. Les descriptions phonétiques les plus poussées se retrouvent dans la plupart des travaux des phonéticiens et pédagogues de la première partie du XXe siècle¹⁸.

- Les liaisons en [z] peuvent être représentées par plusieurs graphèmes:
 - (i) <s> : par exemple *dans un an* sera réalisé [dãzẽnã] ;
 - (ii) <x> : par exemple *de beaux enfants* sera prononcé [dãbozãfã] ;
 - (iii) et <z> : par exemple *chez eux* sera réalisé [ʒezœ].

Les liaisons en [z] sont généralement représentatives d'un pluriel (par ex. *mes amis* [mezami])¹⁹. Martinon (1913) insiste sur cette particularité en rappelant que:

[...] le rôle principal de la liaison ici, celui qu'elle paraît devoir jouer pendant longtemps encore, c'est de marquer le pluriel. Sur ce point, elle ne fléchit guère. C'est pour cela que les articles pluriels, [...], les adjectifs possessifs ou indéfinis, [...], les adjectifs numéraux, [...], se lient sans exception, devant un substantif, bien entendu, même précédé de son adjectif [...]. Ces liaisons sont si nécessaires que le peuple ajoute

¹⁷ Comme nous le verrons, ces paramètres ne permettent toutefois qu'une prédiction approximative. Par exemple, les liaisons variables, sujet de prédilection de la plupart des spécialistes de la liaison indiquent que, dans un même contexte, une consonne peut être réalisée de façon très aléatoire pour un même locuteur (cf. 3.3.3). De la même façon, les liaisons dites *mal-t-à-propos* (cf. 3.3.4) montrent que parfois la consonne réalisée ne correspond pas au schéma établi.

¹⁸ Pour une description détaillée, voir Martinon (1913), Grammont (1914 et 1933), Gougenheim (1935), Fouché (1959), Nyrop (1963) et Delattre (1966).

¹⁹ Cf. p. 24 : remarque de Laks (2005b) à ce propos et citation note de bas de page 3.

volontiers *quatre* à *deux*, *trois*, à *six* [...] : le *Bal des Quat(re) zArts*²⁰.
[...] *Entre quatre z yeux* [...]. (Martinon, 1913 : 375)

- Les liaisons avec [t] peuvent elles aussi présenter des graphies différentes :

(i) <t> (par ex. *tout à fait* [tutafɛ]) peut fonctionner comme marqueur verbal (par ex. *il est arrivé*) signalant le rôle morphologique que la consonne de liaison peut jouer (Morin, 1981). Ce qui a accru la fréquence d'apparition de ce type de liaisons ;

(ii) et <d> : par exemple *grand avenir* [grãtavəniʁ].

Par ailleurs, l'opposition [t] – [d] peut marquer la différence entre le masculin et le féminin (liaison vs enchaînement) comme par exemple dans *grand enfant* [grãtãfã] et *grande enfant* [grãdãfã]. Martinon (1913 : 362) dira de ces deux consonnes (/z, t/) qui lient beaucoup, qu'étant donné le nombre de liaisons qu'elles manifestent, « *c'est ici qu'intervient le goût personnel pour les facultatives* ».

- La consonne de liaison [n] résulte de la présence d'une voyelle nasale de forme V+ /n/ (par ex. *mon* [mõ] et *mon ami* [mõnami]). Martinon (1913) remarque que dans le cas particulier des adjectifs placés devant le substantif, c'est le son du féminin que l'on entend (comparer *lointain* [lwẽtẽ] – *lointaine* [lwẽten] – *lointain ami* [lwẽtenami]).

- Quand aux trois consonnes [p], [g] et [k], elles n'existent que dans des séquences facilement répertoriées en raison de leur nombre très restreint :

(i) [g] et [k], c'est-à-dire les deux prononciations possibles du graphème <g>, ne se retrouvent que dans l'expression *sang impur* [sãkẽpyʁ] de la Marseillaise et dans le mot *long*²¹ (par ex. *long été* [lõgete] vs. [lõkete]) ;

²⁰ A différencier d'avec le *Bal des Ga(rs) d(es)z'arts* dont la prononciation très proche et qui ne correspond aucunement à un velours (cf. 3.3.4), favorise la confusion de sens chez les locuteurs. Le *Bal des Quatre Arts* sera réalisé [baldekãzar] tandis que le *Bal des Gars des Arts* sera réalisé [baldegãzar].

²¹ Certaines études portant sur l'antéposition avancent qu'aujourd'hui la tendance est à la postposition et que seul un groupe restreint d'adjectifs apparaît régulièrement en antéposition (cf. Lyche, 2005). Cependant cette hypothèse est controversée par d'autres spécialistes qui affirment que si certains adjectifs n'apparaissent plus antéposés au nom, cette tendance n'est pas valable pour tous. Nous pensons que s'il est effectivement possible que l'absence de réalisation de la liaison avec un adjectif tel que *long* soit en partie due au fait qu'il est plus souvent postposé qu'antéposé, nous préférons toutefois privilégier l'hypothèse d'une disparition des consonnes de liaison en [k] et [g].

(ii) pour [p], la liaison est limitée aux mots *trop* (par ex. *il l'a trop aimé* [tʁopemɛ]) et *beaucoup* (par ex. *elle s'est beaucoup attachée* [bokuˈpatɑʃe] à *cet animal*).

3.1.2.2 Des changements phonétiques

Lors de l'apparition de ces consonnes de liaison, certains changements phonétiques peuvent avoir lieu, et toucher soit les consonnes, soit les voyelles (Fouché, 1959 ; Martinon, 1913 ; Grammont, 1933 ; Delattre, 1966 ; Selkirk, 1972 et Tranel, 1981).

a) Le (dé)voisement des consonnes : les consonnes de liaison peuvent être touchées par un voisement ou un dévoisement. La fricative finale /s/ se voise (par ex. *gros homme* réalisé [gʁozɔ̃m]), et l'occlusive /d/ connaît un sort inverse, à savoir qu'elle perd son voisement lors d'une liaison (par ex. *grand homme* réalisé [grɑ̃tɔ̃m]²²).

b) La (dé)nasalisation des voyelles nasales : lorsque la liaison a lieu, deux formes phonétiques peuvent apparaître. Soit une dénasalisation de la voyelle (par ex. *bon* [bɔ̃] réalisé [bɔn] dans *bon ami*), phénomène qui touche particulièrement les adjectifs antéposés ; soit la conservation de cette voyelle nasale (par ex. *mon enfant* [mɔ̃nɑ̃fɑ̃]²³). Dans sa thèse portant sur la *Prononciation des consonnes dans le français de Paris aux 13ème et 14ème siècles*, Jong (2006 : 156) fait remarquer que :

[...] la dénasalisation fait l'objet d'un consensus : les voyelles nasales se sont progressivement dénasalisées devant consonne nasale entre le 15ème et le 17ème ou 18ème siècle²⁴, par exemple [bɔ̃nə] > [bɔnə]. Les autres voyelles nasales n'étant plus suivies de consonne nasale (*fond* par

²² L'adjectif *grand* est souvent pris en exemple lorsqu'il s'agit de décrire les variations que peut subir la consonne finale selon les contextes de dérivation (*grande, grandeur*) ou de liaison (*grand (t) escogriffe*). L'analyse de l'évolution de cette consonne finale, touchée par la règle de dévoisement, permet de mieux comprendre ses différentes représentations (cf. 2.1).

²³ Cette généralisation concernant les alternances à voyelles nasales est conforme à la description traditionnelle. Comme nous le verrons dans la section réservée au statut de la consonne de liaison (cf. 4.5), les auteurs ne sont cependant pas tous unanimes : certains préfèrent parler de suppletion et non de processus de (dé)nasalisation. (cf. Morin 1983 ; Tranel, 1990).

²⁴ Il est nécessaire de noter qu'avant ce phénomène massif de dénasalisation, la consonne /n/ a tout d'abord commencé par propager son trait [+nasal] à la voyelle la précédant (XIème – XIIème siècle).

exemple était devenu [fõ(t)], n'ont pas été affectées par la dénasalisation²⁵.

3.1.2.3 Synthèse

Nous récapitulons dans le tableau suivant (Figure 3-1) les différents éléments que nous avons présentés précédemment, soit pour chaque CL : sa phonie, sa propriété de voisement et les différentes graphies possibles.

Phonie	Voisement	Graphies disponibles	Illustration	Remarques
/z/	+	-s, -z, -x	<i>les éléments chez eux délicieux accras suis émerveillée</i>	Souvent marqueur du pluriel
/t/	-	-d, -t	<i>grand ennemi petit aérogare vend une pelle seraient affreux</i>	Souvent marqueur verbal : - singulier - ou pluriel
/n/	-	-(V+) n	<i>aucun ami [+ nas] bon achat [- nas]</i>	Toujours précédé d'une voyelle nasale (qui peut être dénasalisée lors du processus)
/ʀ/	-	-r	<i>premier avion</i>	
/p/	-	-p	<i>trop important</i>	Rare
/g/	+	-g	<i>long été</i>	Rare
/k/	-	-g		

Figure 3-1 Synthèse sur les caractéristiques de chaque CL

3.1.3 Une structure syllabique

Il est généralement admis que le français présente une forte tendance à la syllabation ouverte (cf. 2.1). Cette tendance est soulignée par Laks (2005b) qui, à la suite de Martinet (1964 : 328), rappelle que la liaison :

[...] correspond à un universel qui privilégie la création et la stabilisation de suites consonnes voyelles. Comme on le voit, la liaison avec resyllabation de la consonne finale à l'initiale du mot suivant peut souvent correspondre à un optimum d'équilibre entre la dynamique

²⁵ Cf. Morin 1994 et 2003.

d'enchaînement généralisé et la dynamique de syllabation ouverte.
(Laks, 2005b : 162)

La liaison serait donc un moyen de parvenir à cette alternance CV (consonne voyelle), considérée comme optimale, lorsque le contexte phonétique s'y prête. Pour comprendre les stratégies mises en place, il est nécessaire de revenir sur le concept de syllabe. Pour cela, sans proposer une analyse exhaustive de la syllabe, nous adopterons une définition en nous appuyant sur les travaux d'Encrevé (1988 : 138)²⁶.

3.1.3.1 Représentation formelle de la syllabe

La syllabe (σ) est généralement considérée comme une structure binaire²⁷ branchante qui se subdivise en une attaque (A), élément consonantique initial et une rime (R). Cette dernière est minimalement formée par un noyau vocalique (N) et éventuellement par une coda (C) (notée entre parenthèses dans le schéma ci-dessous du fait même de son optionalité) :

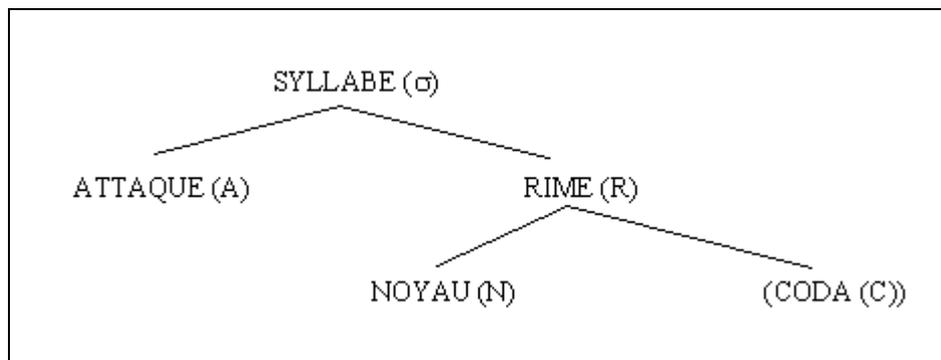


Figure 3-2 Structure syllabique minimale

Ce modèle, que l'on dit universel, considère que le noyau vocalique, seul élément obligatoire, constitue le pilier de la structure autour duquel s'agencent les segments consonantiques facultatifs.

²⁶ Encrevé (1988) s'appuie principalement sur Kaye et Lowenstam (1984) ne retenant de leurs travaux que les aspects lui paraissant en jeu dans le phénomène étudié (Encrevé, 1988 : 243, *note 2*) et analysera la structure syllabique minimale sur les modèles de Pike et Pike (1947) et Kurylowicz (1948).

²⁷ Des conceptions ternaires de la syllabe CVC existent aussi (voir notamment Clements, 1983).

L'attaque peut être phonétiquement vide, mais elle est structurellement présente, en creux, en tant que position d'accueil (voir le phénomène de la liaison en français par exemple, [...], ou les phénomènes d'insertion d'une consonne par défaut, type « coup de glotte » dans les attaques vides en allemand ou *hamza* en arabe). La coda en revanche peut être totalement absente. Toute syllabe est donc obligatoirement constituée d'une attaque et d'une rime, dont le noyau est obligatoirement non-vide. (Labrune, 2005 : 101)

Dans un cadre multidimensionnel, Encrevé (1988) postule que toutes les codas sont flottantes, aussi bien les consonnes finales fixes que les consonnes de liaison :

En français, à la finale de mots, un constituant syllabique *coda* est *toujours flottant*, qu'il corresponde à une CF ou à une CL. Ce fait de structure se traduit directement dans le fait empiriquement constaté de l'enchaînement. Il suggère une hypothèse beaucoup plus générale : *en français, tout autosegment syllabique coda est flottant*. (Encrevé, 1988 : 175)

Ces autosegments de la ligne segmentale ne sont donc pas ancrés dans le squelette (cf. 4.2).

3.1.3.2 Un enchaînement ou pas

Dans une perspective multilinéaire²⁸, la consonne de liaison est vue comme un segment flottant appartenant lexicalement au mot au niveau sous-jacent, et qui possède, au même titre que les consonnes fixes, une position dans la syllabe bien qu'elle n'y soit pas associée. Pour qu'elle soit interprétée segmentalement, elle se doit de répondre à la condition paramétrique suivante :

[...] une consonne finale flottante ne peut s'ancrer dans le squelette que si le mot suivant dans la chaîne parlée commence par une attaque nulle. (Encrevé, 1988 : 179)

Encrevé (1988) décrit ainsi la liaison comme l'association d'une consonne flottante dans une position d'attaque flottante nulle²⁹. Lorsque la consonne de liaison est associée

²⁸ Nous proposons en 4.2 (pp. 121 et *sq.*), une présentation des formalismes proposés dans le cadre de la phonologie multilinéaire.

²⁹ Encrevé (1988), à la suite de Vergnaud (1982), préférera l'utilisation le terme *nulle* à celui de *vide*. Dans le cas des mots « à initiale vocalique » pour lesquels *l'attaque* n'a aucune correspondance sur la ligne segmentale (non ancrée dans le squelette et vide de tout contenu segmental dans la forme lexicale du mot), Encrevé estime qu'« en termes d'intuition ordinaire du locuteur, aucune consonne initiale

à l'attaque vide du mot suivant, elle peut être réalisée, annulant par là même la coda de la première unité lexicale.

Illustrons ce processus avec un exemple repris à Encrevé (1988 : 177 et *sq.*) : la phrase *j'avais un rêve* [ʒavezɛ̃rev] dans laquelle [z] est la consonne de liaison.

Dans le premier schéma (Figure 3-3), correspondant à un contexte potentiel de liaison, la consonne de liaison [z] n'est pas rattachée au squelette, elle est flottante tout comme la coda (C) lui correspondant. De la même façon, l'attaque de la seconde unité lexicale, constituant syllabique initial de celle-ci, est flottante et nulle.

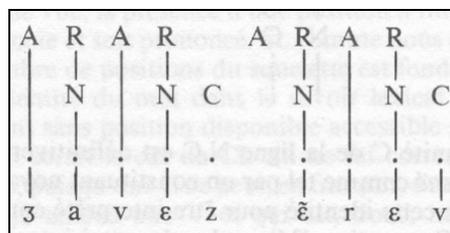


Figure 3-3 Représentation où la consonne flottante appartient au premier mot (Encrevé 1988 : 172)

Dans le schéma suivant (Figure 3-4), on peut observer l'association de la consonne flottante [z] à l'attaque flottante qui lui succède. La consonne de liaison se retrouve alors interprétée comme attaque syllabique du second mot.

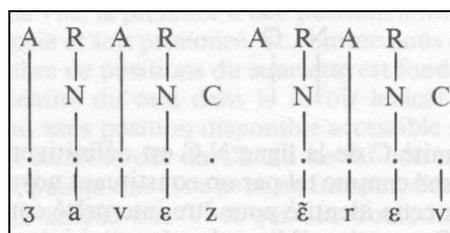


Figure 3-4 Association du segment flottant à l'attaque flottante qui le suit (Encrevé, 1988 : 177)

potentielle n'est comptée au nombre des segments du mot ». Cette attaque ne peut donc être simplement *vide*, elle est également *nulle*.

Nous savons à présent quels éléments sont constitutifs du phénomène de la liaison et nous sommes désormais à même de déterminer les deux « moments » phonétiques distincts qui permettent de le définir :

1) l'apparition à la finale d'un mot d'une consonne latente, de liaison, phénomène qui se produit devant initiale vocalique et ce, dans des conditions syntaxiques précises, que nous explorerons plus loin dans notre recherche (*cf.* 3.2, pp. 57 et *sq.*) ;

2) la nouvelle syllabation qui s'ensuit où, dans le cadre général de l'enchaînement caractéristique du français, la CL s'entend comme formant syllabe avec l'initiale vocalique du mot suivant.

En adoptant le modèle d'Encrevé (1988), nous admettons l'existence d'un *non-enchaînement*, propos central de son ouvrage. A sa suite, nous considérons que le non-enchaînement correspond à la non-association syllabique de la consonne finale, réalisée, à la voyelle initiale du mot suivant. Ainsi, Encrevé (1988 : 29) :

Nous posons donc qu'enchaînement et non-enchaînement peuvent se constater en français dans la chaîne parlée pour toutes les consonnes placées devant un mot à initiale vocalique, que ces consonnes soient toujours prononcées (Consonnes Fixes, dorénavant notées CF) ou qu'elles ne s'entendent que dans ce contexte particulier (CL).

3.1.4 Le hiatus

Traditionnellement, la capacité qu'a le français de lier les termes les uns à la suite des autres, en enchaînant, en liaisonnant, serait due à une intolérance des séquences d'enchaînement vocalique (appelé *hiatus*). La liaison n'a en effet lieu que dans une séquence de deux mots (M1 et M2) où M1 se termine par une voyelle et M2 commence par une voyelle³⁰.

³⁰ Encrevé (1988 : 58) fait état de très rares cas d'emploi (2 au total) d'une CL dans un contexte phonétique où elle est, par définition, impossible : devant un mot à initiale consonantique (par ex. *quand monsieur* réalisé [ka.tmə.sjœ]). Dans cet exemple, la consonne de liaison [t] est syllabée à l'attaque de la première syllabe du mot suivant, comportant déjà une initiale consonantique. Nous ne tiendrons cependant pas compte de ces observations. Le nombre de ces emplois est, comme Encrevé le dit lui-même, (i) extrêmement faible et donc non représentatif comparé à l'ensemble des 7 842 CL qu'il a examinées et qui sont réalisées devant voyelle et (ii) traduit par conséquent un phénomène agrammatical.

En français, nous disons *deux pommes*, mais *deux_oranges* ; *les coutumes* mais *les_habitudes* ; *le vieux chêne*, mais *le vieil_arbre* ; *l'an nouveau*, mais *le nouvel_an* ; *le marteau*, mais *martel_en tête* ; *il dit*, mais *dit-il* ; *il va*, mais *va-t-il* ; *mon_amie*, pour *ma amie*. Le génie de notre langue s'oppose à l'hiatus. Le rythme phonétique du français est tatata, c'est-à-dire dans l'alternance des consonnes et des voyelles. Aussi rejette-t-il l'accumulation des consonnes et la contiguïté de deux voyelles. (Tribouillard, 1998 : 17)

Le hiatus correspond à la rencontre de deux voyelles appartenant à des syllabes différentes. Il peut se produire dans deux cas de figure différents :

(i) **hiatus interne** : soit à l'intérieur d'un même mot comme par exemple dans *créole* [kreɔl] ou encore dans *épiploon* [epiplɔ̃] ;

(ii) **hiatus transitoire** : soit dans la succession de deux mots (par ex. *il va à l'école*, *point final* ! /va.a./).

Le terme *hiatus*³¹ est défini par le *Dictionnaire de l'Académie française* (1762, 4^{ème} édition)³² comme :

[...] emprunté du Latin, pour exprimer une sorte de baillement, qui fait un méchant effet dans la Poésie. Dans les vers, il faut éviter les hiatus. Un hiatus qui blesse horriblement l'oreille. L'hiatus est proprement entre une voyelle qui finit un mot, et celle qui en commence un autre, et où il ne se fait point d'élision.

Cette définition qui, en premier lieu ne concerne que la récitation des vers en poésie, s'est vue imposée à l'ensemble des registres de la langue française et principalement aux règles établies pour la liaison avec le triomphe du *bon usage* au XVIII^e siècle (cf. 2.2., pp. 24 et sq.). Le hiatus est alors considéré comme une faute sans rémission dans la langue parlée. Aujourd'hui encore, les linguistes ne parviennent pas à s'accorder pour déterminer si le hiatus relève plus d'une norme imposée, dont l'influence s'est

³¹ Du latin *hiatus*, ouverture, de *hio*, béer, *hiatus oris*, bouche ouverte, parce que la bouche s'ouvre dans le hiatus.

³² Version révisée en 1986 (9^{ème} édition) : « [...] Emprunté du latin *hiatus*, « action d'ouvrir, ouverture » puis en rhétorique, « rencontre de deux voyelles dont la prononciation conduit à maintenir la bouche ouverte ». [...] Rencontre de deux voyelles appartenant à des syllabes différentes, se produisant soit dans la succession de deux mots, soit à l'intérieur d'un mot. Il y a un hiatus dans « tu as », « tu es ». Il y a un hiatus interne dans « boa », « séisme ». Ce hiatus blesse l'oreille. Le hiatus d'un mot à un autre a été proscrit de notre poésie par Malherbe ».

renforcée au fil des siècles ou d'une tendance qui a toujours eu cours, permettant d'arguer que le français prohibe réellement le hiatus :

L'exégèse du vers classique a rationalisé certains compromis sociaux en détournant et en amplifiant certains concepts de la métrique du latin et du grec ancien. Elle a en particulier insisté sur le rôle de la « liaison » comme un des moyens pour assurer l'absence d'hiatus entre deux mots, conformément aux canons esthétiques du vers (*cf.* l'analyse de la fiction graphique de Cornulier 1995, et les études historiques de Morin 2000 et 2005). (Morin, 2005 : 8)

Martinon (1913) fait remarquer que de nombreuses consonnes ont, au cours du XVIIIe siècle, perdu leur oralité. Si certaines ont été conservées, c'est, écrira-t-il (1914 : 359), « *en principe dans une intention d'harmonie, et notamment pour éviter les hiatus* ». Cependant, certains contextes ne bénéficient aucunement de l'appui de la liaison pour éviter le hiatus (par ex. *tu as_encore séché l'école !* [ty.a.ã.kɔʀ] ; *il va_à Alès en voiture* [il.va.a.ales.]), et la structure syllabique reste intacte contrairement à la resyllabation. Qu'en-est-il réellement ? Dans cette section, nous nous attacherons à confronter les différents points de vue.

Lorsqu'il est question du hiatus dans le cas de la liaison, on trouve des descriptions très fines du phénomène mais qui montrent que sa prohibition n'est pas totale :

L'hiatus est quelquefois doux, quelquefois dur ; [...] non seulement l'hiatus est quelquefois permis, mais il est souvent agréable. C'est au sentiment de le choisir, c'est à l'oreille de marquer sa place. [...] L'hiatus sera donc mélodieux dans la liaison des mots, car il est égal pour l'oreille que les voyelles se succèdent dans un seul mot, ou d'un mot à l'autre. [...] Nous éprouvons cependant qu'il y a des voyelles dont l'assemblage déplaît : *a-w, o-i, a-an, o-un*, [...] et l'on en trouve la cause physique dans le jeu même de l'organe. Mais deux voyelles dont les sons se modifient par des mouvements que l'organe exécute facilement, comme dans *Clio, Danaé*, non seulement se succèdent sans dureté, mais avec beaucoup de douceur. L'hiatus d'une voyelle avec elle-même est toujours dur à l'oreille [...] La règle qui défend l'hiatus est donc une règle capricieuse, et aussi peu en accord avec elle-même qu'avec l'oreille, qu'elle prive d'une infinité de douces liaisons. (Laveaux, 1847 : 345)

Chez Grammont (1914 et 1933), la liaison est traitée par la phonétique impressive, qui propose une analyse des sentiments et impressions que provoquent les sons chez les

auditeurs. Il établit une distinction entre la liaison vocalique, où il y a absence de consonne entre les deux voyelles (ce qui correspond au hiatus), et la liaison consonantique (qui est la liaison proprement dite) qui a toujours lieu dans un groupe rythmique. La définition qu'il propose de la liaison est ainsi indissociable de celle du hiatus : si dans une même phrase, deux mots se suivent, que le second commence par une voyelle et que le premier se termine par une consonne qui ne se prononce pas dans toutes les positions, deux cas de figure sont possibles : soit une liaison si cette consonne est prononcée (liaison consonantique), soit un hiatus si elle ne l'est pas (liaison vocalique).

Il pose comme principe général que **la liaison consonantique** est dépendante du groupe rythmique. On lie toujours à l'intérieur d'un groupe rythmique, on ne lie pas entre deux groupes rythmiques. Autrement dit, on lie d'une syllabe accentuée sur la suivante, on ne lie pas depuis une syllabe inaccentuée. Le plus souvent, l'application de cette règle permet de distinguer entre deux interprétations sémantiques. Ainsi, dans *savant aveugle*, la liaison sera réalisée lorsque *aveugle* sera substantif (*savant*, adjectif est ainsi préposé) mais non lorsqu'il sera adjectif et donc postposé au substantif (*savant*).

A propos de **la liaison vocalique**, le hiatus, Grammont procède à une description d'ordre plus articulatoire. Il pose que cette liaison est variable selon les langues et que l'on observe (i) soit un coup de glotte (donc une consonne laryngale sourde) avant la deuxième voyelle, (ii) soit une vibration uniforme de la glotte qui donne lieu à une liaison vocalique « vraie ». Selon lui, en français, la glotte ne cesse pas de vibrer, il n'y a donc aucune interruption mais plutôt un prolongement articulatoire. Il en conclut qu'en français il y a toujours liaison, soit consonantique, soit vocalique. Les liaisons vocaliques comportent pour la plupart une « *modulation de timbre et de hauteur qui les rend infiniment douces et agréables* » (Grammont, 1914 : 136). Cette « douceur » serait responsable du fait que certaines liaisons consonantiques sont de moins en moins réalisées puisqu'il n'en « *résulte aucun inconvénient ni aucune difficulté de prononciation* » (1914 : 136).³³

³³ De la même façon, Delattre (1955) indique que la liaison est plus souvent réalisée quand les voyelles sont de deux timbres différents. La liaison s'opère plus facilement lorsqu'il s'agit d'éviter un hiatus entre deux voyelles de timbre différent.

La plupart des auteurs classiques rejoignent plus ou moins cette position selon laquelle la liaison, lorsqu'elle est réalisée, permet d'éviter le hiatus mais qu'elle ne constitue pas une stratégie de « réparation du vide entre deux voyelles » au profit d'une syllabation ouverte³⁴. Encrevé (1988), en particulier, déplore la surdité des orthoépistes face à la réalité des usages, à savoir que la prohibition des hiatus est réservée à des usages particuliers (par ex. les professionnels de la parole publique, les acteurs, *etc.*) :

Mais aujourd'hui comme autrefois le dire de l'abbé Olivet reste vrai : « La conversation des honnêtes gens est pleine d'hiatus volontaires qui sont tellement autorisés par l'usage, que, si l'on parlait autrement, cela serait d'un pédant ou d'un provincial ». (Encrevé, 1988 : 260)

Il rappelle que ces propos sont attestés par de nombreux auteurs. Martinon (1913 : 391-392) par exemple, reprend la position d'Emile Littré (1866) lequel, bien qu'il admette que les liaisons sont plus souvent réalisées qu'aux XVIIIe et XIXe siècles, indique qu'il ne faut pas sacrifier la tradition des usages mais la suivre : la liaison ne se fait que dans un certain nombre de cas prédéterminés. De la même façon, Clédat (1917) précise que « *les liaisons empêchent des hiatus, mais elles ne sont pas, comme on le croit souvent, destinées à les éviter, et elles ne reposent pas sur une raison euphonique.* »³⁵

Ågren (1973) a soumis à un examen soigneux l'hypothèse selon laquelle la liaison est plus souvent réalisée lorsqu'il y a hiatus, en la testant aux occurrences de *noms pluriels + adjectif ou participe passé* présentes dans son corpus³⁶. Il répartit les noms pluriels en trois groupes : (i) les noms se terminant par une voyelle ; (ii) les noms se terminant par une consonne et (iii) les noms se terminant par deux consonnes.

³⁴ De la même façon, Morin (2005 : 17-18), dans son article consacré exclusivement à une réflexion historique autour de la problématique du hiatus, fait référence aux alternances Consonne Zéro et indique qu'elles résultent uniquement de la chute de consonne devant consonne : « *l'évolution générale du français révèle une absence totale d'effets dus à la contrainte anti-hiatus [...] sur la liaison. Au XVIème siècle encore, les grammairiens conservateurs décrivent des usages où la prononciation des obstruents médiévales s'est conservée à la pause et devant voyelle, comme on peut voir dans ce beau morceau d'anthologie tiré des Hypomneses de Henri Estienne (1582 : 94-95), que Thurot avait choisi pour illustrer la troncation des consonnes au XVIème siècle* ».

³⁵ A ce propos, voir Encrevé (1988 : 260).

³⁶ Pour une lecture détaillée de l'étude, cf. Ågren (1973 : 128-130).

Il constate alors que les différences de réalisation de la consonne de liaison entre les trois groupes sont infimes et fonctionnent même plutôt de façon inverse aux prévisions de Delattre (1955). Les résultats obtenus sont repris et résumés dans le tableau suivant :

	Nombre d'occurrences	Nombre de liaisons réalisées	Nombre de liaisons non réalisées	% de réalisation
N terminé par 1 voyelle	281	94	187	33,4
N terminé par 1 consonne	199	70	129	35,1
N terminé par 2 consonnes ³⁷	53	22	31	41,5

Figure 3-5 La liaison et le hiatus (Ågren, 1973)

Ågren en conclut, sans oser « *affirmer que la liaison se [fait] plus fréquemment après consonne qu'après voyelle* » (1973 : 128), que la liaison est aussi souvent pratiquée après consonne qu'après voyelle, les enchaînements consonantique et vocalique ayant les mêmes effets sur la nécessité de lier. Il ajoute qu'en cas de liaison plurielle en /z/ après deux consonnes, le locuteur peut avoir recours au schwa pour faciliter la prononciation de la consonne de liaison.

Comme le rappelle Morin (2005 : 8), les premières analyses structurales ont fait peu de cas du hiatus (par ex. Gougenheim, 1935). Elles procèdent à une analyse de la consonne de liaison par supplétion³⁸. *De facto*, le problème du hiatus n'en est plus un.

Cependant, le développement de nouveaux modèles théoriques (par ex. Théorie de l'Optimalité³⁹, cf. 4.3, pp. 128 et sq.), redéfinit les consonnes de liaison comme étant des « consonnes latentes », c'est-à-dire qui « *apparaissent seulement devant voyelle dans certains contextes syntaxiques et stylistiques* » (Tranel, 2000 : 47-48). Pour souligner que certes le hiatus est rare, mais n'en demeure pas moins possible, l'OT définit une contrainte, interdisant le choc vocalique : *VV. Et il apparaît qu'en règle générale, cette contrainte est respectée.

³⁷ Parmi les 53 occurrences de *noms pluriels terminés par deux consonnes*, Ågren en relève 16 du type [tR], [vR], [Rs], entre autres, et pour lesquels les locuteurs ont inséré un *e* instable (schwa) (cf. Ågren, 1973 : 129, note 21).

³⁸ Cf. 4.5.

³⁹ Désormais OT.

Steriade (1999), toujours dans le cadre de l'OT, met en relation morphologie et phonologie en étudiant la liaison après un adjectif préposé au nom. Elle tente d'expliquer pourquoi c'est, en général, la forme au féminin d'un adjectif qui est utilisée dans le cas de liaison au masculin : *beau navire* [bo], *bel avion* [bɛl], *belle fleur* [bɛl]. Sans entrer dans les détails de son argumentation, rappelons qu'elle considère que les formes lexicales proches, sont organisées en paradigmes. Dans le cas de la liaison, la formation de la forme liaisonnée (*petit* /pətiti/) s'intègre au paradigme lexical du mot en question. La consonne de liaison est issue de la forme au féminin de l'adjectif ; elle est utilisée aussi au masculin pour éviter le hiatus : « *The only solution that resolves hiatus without resorting to the creation of a phonological novel form* » (Steriade, 1999 : 252).

La contrainte *VV est aussi centrale dans l'étude de Tranel (2000), qui stipule que deux voyelles ne doivent pas se succéder dans deux syllabes différentes ou à la frontière de deux mots. Cette contrainte est en concurrence avec d'autres, telles que la contrainte de fidélité entre l'input et l'output (ce qui revient à éviter les insertions et les suppressions de phonèmes).

Toutefois, les contextes envisagés par les auteurs travaillant dans le cadre de l'OT sont singulièrement restreints. Le bilan critique a été effectué par Morin (2005) et est résumé ci-dessous :

[...] la contrainte anti-hiatus (ou la contrainte anti-attaque vide) apparaît comme un simple moyen mécanique d'assurer que les consonnes de liaison apparaissent devant un mot à initiale vocalique. En se dispensant de faire l'analyse des « contextes de liaison », elles [les nouvelles analyses phonologiques] sont incapables d'expliquer pourquoi les mêmes contraintes sont inopérantes ailleurs, et par là-même d'apprécier la nature des différents types de liaison. (Morin, 2005 : 16)

Citons enfin, les travaux de Bonami et Boyé (2003), qui, dans une analyse en termes de « Compétition entre Phonologie et Syntaxe », expliquent que la généralisation descriptive selon laquelle l'allomorphie est motivée par une raison phonologique (évitement du hiatus) est erronée et ne peut s'appliquer qu'à un contexte syntaxique spécifique : *déterminant et/ou adjectif prénominal et le mot qui suit*. Ils concluent sur le fait que la contrainte sur les formes de liaison est loin d'être une pure contrainte phonologique locale.

Au bilan, il apparaît que :

1) soit le hiatus n'est pas un problème, car on effectue une analyse en termes de supplétion (*cf.* les analyses structurales) ;

2) soit le hiatus est un problème. Les explications qui ont été apportées ne sont pas convaincantes, étant donné le nombre restreint de contextes qui a été examiné.

Il apparaît en outre que si le rôle de la liaison était d'empêcher les hiatus, alors on ne devrait pas pouvoir formuler des expressions telles que :

(1) *Il avait_encore bu : sans liaison ; cas de liaison variable*

(2) *L'avion_américain est tombé en panne : sans liaison ; cas de liaison erratique*

En fait, comme l'illustrent les exemples ci-dessus, le hiatus en français est possible. Or, dans la littérature, le problème n'a jamais été posé dans sa généralité :

1) en effet, parce qu'on n'a pas vu le lien entre hiatus et liaisons variables, la question du hiatus ne se pose véritablement que dans le cas des liaisons variables ;

2) et les arguments en faveur d'une stratégie d'évitement du hiatus proviennent de l'examen des contextes de liaisons catégoriques.

Enfin, le traitement du hiatus dans la littérature fait apparaître les limites d'un modèle qui s'appuierait uniquement sur la phonologie pour expliquer le phénomène de la liaison (Bonami et Boyé, 2003). Nous dressons le bilan de ces limites dans la section suivante.

3.1.5 Les carences d'une explication phonologique

La liaison s'effectue au contact de deux mots, à partir du moment où le premier mot contient une consonne flottante en finale et où le second mot est à initiale vocalique.

Les exemples suivants proposent un ensemble de phrases contenant chacune un contexte potentiel de liaison. Dans les exemples (3) et (4), la liaison est toujours

réalisée. Dans les exemples (5) et (6), elle est variable et dans l'exemple (7), elle n'est jamais réalisée.

- (3) *Certains_enfants coûtent cher à leurs parents* [sɛʁtɛ̃zãfã] / *[sɛʁtɛ̃ãfa]
- (4) *J'en_ai marre* [ʒãnɛmar(ə)] / *[ʒãɛmar(ə)]
- (5) *Ce serait_inconvenant* [s(ə)s(ə)ʁɛtɛ̃kɔ̃v(ə)nã] / [s(ə)s(ə)ʁɛ̃kɔ̃v(ə)nã]
- (6) *Il a pris des livres_à la bibli* [delivʁ(ə)alabibli] / [delivʁəzalabibli]
- (7) *Il a adopté un chien_abandonné* [ɛ̃ʃjɛabãdone] / *[ɛ̃ʃjɛnabãdone]

Ces exemples montrent que la liaison est un phénomène variable pour lequel la condition phonologique est insuffisante pour rendre compte de la complexité du phénomène. Si une analyse phonologique de la liaison permet d'en décrire les mécanismes, elle n'explique cependant pas ce qui la caractérise le plus : sa variation. C'est d'ailleurs ce que mettent en avant respectivement Encrevé (1988) et Laks (2005a) :

La satisfaction de la Condition [paramétrique d'ancrage des consonnes finales flottantes, cf. 44 et sq.] est nécessaire à la réalisation de la liaison, mais elle n'est pas suffisante. La liaison n'a lieu effectivement que si, en outre, les conditions non phonétiques (syntaxiques, stylistiques, etc.) [...] sont satisfaites. (Encrevé, 1988 : 181)

Une analyse strictement limitée au plan phonético-phonologique ne parvient pas mieux à expliquer la variabilité interne du phénomène, caractéristique pourtant indépassable de la liaison en français. En effet, les variables sociolinguistiques externes (diachronique, diatopique, diaphrastique) étant tenues constantes, la liaison reste pourtant un phénomène intrinsèquement variable : dans le même discours, le même locuteur, au même moment peut faire la liaison puis, après une rupture intonative et une reprise immédiate, ne plus la faire. [...]. On comprend immédiatement qu'une analyse phonologique stricte ne pourra expliquer une telle variation intrinsèque [...]. (Laks, 2005a : 159)

La phonologie permet de circonscrire les contextes de liaisons possibles, un contexte correspondant à la présence d'une consonne flottante finale devant une attaque vide. Mais seule, elle ne peut expliquer le fait que certains des contextes manifestent une liaison dite catégorique, que d'autres manifestent une liaison dite variable et que d'autres encore ne manifestent aucune liaison. Comme le soulignent Bonami et Boyé (2003 : 42), « *Si le conditionnement était purement phonologique, la liaison devrait être partout obligatoire en français* ».

Laks (2005a) va plus loin dans le raisonnement, arguant que « *le plan phonologique et l'analyse segmentale contextuelle ne définissent qu'une virtualité de liaison dont l'actualisation est réglée à un autre niveau* » (2005a : 157), et que « *la phonologie syllabique et la présence d'une attaque vide n'en définissant que la virtualité [de la liaison]* » (2005a : 157). Selon lui, la différence de fréquence de la liaison est motivée par l'incapacité du contexte phonologique à expliquer la tripartition des liaisons (variables, catégoriques, erratiques). Il prend comme exemple la consonne de liaison /z/ et remarque qu'un même segment peut tout à la fois se trouver en contexte de liaison variable, de liaison catégorique et de liaison erratique.

Si l'on regarde les exemples que nous avons présentés en début de cette section, on constate qu'effectivement un même segment peut intégrer différents contextes. Ainsi, les exemples (3) et (6) présentent une même consonne flottante : /z/. Or, si la liaison est catégorique en (3) entre *certains* et *enfants*, elle n'est que variable en (6) entre *livres* et *à*. De la même façon, dans les exemples (4) et (7), le segment /n/ apparaît tout à la fois dans une liaison dite catégorique (4) et dans une liaison erratique (7).

Aussi, pour rendre compte et expliquer cette variation, faut-il :

[...] prendre en compte un grand nombre de facteurs et de niveaux d'analyse (rythmique, prosodie, sémantique discursive et pragmatique), non au titre d'un contexte externe influant sur un mécanisme phonologique constant, mais comme véritable moteur et explication de la variation observée. (Laks, 2005a : 159)

Considérer les différents niveaux d'analyses linguistiques constitue le second point. La plupart des travaux traditionnels portant sur des niveaux autres que phonologique ne considèrent généralement que la liaison variable. Lorsque Delattre (1955) se propose d'expliquer « *les facteurs de la liaison* », il faut entendre « *les facteurs influençant la*

fréquence de la liaison variable ». De la même façon, l'ouvrage de Ågren (1973), dans lequel sont testés un certain nombre de facteurs linguistiques, ne porte que sur les liaisons facultatives. Quelques travaux étendent toutefois l'étude de ces facteurs à l'ensemble des contextes. Fougeron *et al.* (2001b), par exemple, ont choisi de tester l'influence des facteurs stylistiques, syntaxiques et lexicaux sur l'ensemble des contextes de liaisons « possibles », que la liaison soit effective ou non.

Aussi est-il nécessaire d'inclure dans le champ des investigations (i) tous les contextes de liaisons définis (variables, catégoriques et erratiques) sur lesquels (ii) pourront être testés les facteurs décrits comme déterminants pour la réalisation de la liaison.

Ces facteurs que nous proposons de passer en revue dans les sections suivantes sont :

- les facteurs morphosyntaxiques, à savoir la cohésion des éléments à lier et le rôle fonctionnel de certaines consonnes de liaison ;
- les motivations socio-stylistiques ;
- et les facteurs lexicaux, à savoir la longueur et la fréquence des occurrences.

3.2 Morpho-Syntaxique

L'analyse de la dimension morphosyntaxique comme facteur influençant la réalisation de la liaison compte parmi les thèmes les plus étudiés. On peut différencier trois ensembles de travaux :

(i) ceux qui se sont attachés à son rôle fonctionnel, essentiellement celui de marqueur du pluriel⁴⁰ ; mais parmi /t/ et /z/, c'est la liaison en /z/ qui a été la plus étudiée⁴¹ ;

⁴⁰ Soit son rôle sémantique. On pourrait y apporter les différentes notations éparses concernant d'autres aspects structuraux et par suite sémantique. Par exemple, Gougenheim (1938 : 60) note qu'à l'intérieur du groupe nominal, la liaison construit deux sens différents selon qu'elle est ou non réalisée : « *tandis qu'elle ne se fait pas d'ordinaire dans le groupe substantif + adjectif, elle a lieu dans le groupe adjectif + substantif.* » On pourra donc distinguer l'expression *un savant anglais*, dans laquelle la liaison non réalisée indique que *savant* est substantif, *anglais* est adjectif de l'expression *un savant Anglais*, dans laquelle la liaison est réalisée en -t, *savant* étant adjectif et *Anglais* substantif (cf. p. 48).

(ii) ceux qui se sont attachés à la spécification de contextes, dans lesquels la liaison était catégorique, variable ou erratique. Il s'agit essentiellement des travaux des descriptivistes ;

(iii) enfin, ceux qui se sont attachés à décrire les contraintes syntaxiques pesant sur le phénomène. On peut regrouper ici les travaux plus récents, s'inscrivant dans une approche formelle.

Les travaux notés sous (ii) et (iii) ne sont pas sans rapport :

- ils s'accordent tous sur le fait que la syntaxe est une dimension déterminante dans la réalisation de la liaison et, généralement, conjointement avec d'autres dimensions ;

- la typologie des contextes mise au point par les travaux des descriptivistes a constitué la base empirique de beaucoup de travaux formels. De plus, il ne s'est pas agi de revenir sur la typologie proprement dite mais plutôt sur son contenu, voire sur les moyens de l'exprimer.

En effet, ce qui est entendu par syntaxe n'est pas défini, selon les approches descriptivistes ou formalistes, avec les mêmes moyens :

1. Certains travaux font de la description syntaxique (voire réduisent la description syntaxique à) une suite linéaire de catégories ; par exemple, est dite syntaxique, une séquence du type « Déterminant + Nom », « Pronom + Verbe », *etc.* La caractéristique de ces travaux est que, le plus souvent, on produit des descriptions à partir de ces suites linéaires de catégories. On trouve cette approche dans les travaux des descriptivistes et dans des travaux plus récents dont le thème central n'est pas la syntaxe ;

2. D'autres travaux fournissent des analyses en termes syntagmatiques ; ils font appel aux notions de groupes / syntagmes, d'organisation interne de ces syntagmes, et éventuellement de relations entre ces syntagmes. Ces travaux s'inscrivent le plus souvent dans un cadre de linguistique formelle, et ont pu aussi mener une analyse bivoire tridimensionnelle, en rapprochant les aspects syntaxiques d'autres dimensions⁴².

⁴¹ Cf. note de bas de page 3 (p. 24) concernant la résistance du *s* face à l'amuissement général des consonnes finales.

⁴² On peut citer tout d'abord les analyses syntaxiques transformationnelles, s'inscrivant dans un cadre générativiste. Les travaux de Selkirk (1972 et 1974) en sont une très bonne illustration : on postule une règle syntaxique permettant d'appréhender le phénomène dans sa généralité ; cependant, confronté à différentes difficultés, son travail se déporte sur les relations entre syntaxe et phonologie, pour essayer de

Dans cette partie, nous développerons essentiellement les approches notées sous 1., soit les approches des descriptivistes, qui font de la description syntaxique une suite linéaire de catégories⁴³. Nous commencerons cependant par faire état des résultats qui ont trait à la liaison en /z/, soit le pluriel.

3.2.1 La liaison en /z/ : le pluriel

Il est généralement observé qu'un même contexte syntaxique peut présenter des fréquences de liaisons différentes en fonction du statut morphologique des éléments (cf. 3.1.2.1, p. 39, citation de Martinon, 1913). En 1914, Grammont, à l'occasion d'une discussion sur la conservation de la liaison, note le rôle fonctionnel qu'elle peut parfois remplir :

Les nouvelles générations tendent à ne garder que les liaisons qui jouent un rôle grammatical, particulièrement celles qui servent à distinguer le singulier du pluriel. (Grammont, 1914 : 130)

Il se trouve qu'une construction de type « Nom + Adjectif » peut réaliser la liaison en /z/ (Gougenheim, 1938 : 60) : « *La liaison en z peut avoir une valeur morphologique, même après r ; un cours élémentaire, sans z, des cours élémentaires, avec z.* » Mais cette consonne de liaison ne peut se réaliser phonétiquement et morphologiquement que lorsqu'elle fonctionne comme un préfixe pluriel de l'adjectif.

rendre compte des variations stylistiques des emplois de la liaison variable. Avec les travaux de De Jong (1994), par exemple, on observe un rapprochement des analyses générativistes avec une approche multilinéaire ; plus récemment, les travaux de Bonami, Boyé et Tseng (2003 et 2005) cherchent à intégrer toutes les dimensions du phénomène, dont la dimension syntaxique, dans le cadre d'une grammaire HPSG. Quels que soient ces travaux, la partition des liaisons proposée par les descriptivistes n'est jamais remise en question. Elle sert de point d'ancrage à toutes les analyses.

⁴³ Le détail des propositions formelles et de leurs critiques n'est pas ici essentiel à notre propos. D'une part, de nombreux travaux en ont fait l'exposé et la critique (on se reportera utilement à Morin et Kaye, 1982 ; Turcsan, 2005 et Dugua, 2006). D'autre part, notre travail est fondé sur l'analyse de données extraites du projet PFC : données qui ont été extraites selon une certaine méthodologie (cf. 5.1, pp. 145 et sq.) et pour lesquelles nous ne disposons pas d'informations permettant de tester en détail ces propositions formelles. Par contre, nous disposons de beaucoup plus d'éléments permettant de tester les propositions des descriptivistes. Comme nous le verrons, l'un des résultats essentiels auxquels ils arrivent est la tripartition des types de liaisons, tripartition dont nous avons souligné dans la note précédente (42) qu'elle n'avait jamais été remise en question par les formalistes. C'est pourquoi nous consacrons l'essentiel de cette présentation aux travaux des descriptivistes.

De la même façon, Delattre (1947) indique que la liaison se fait davantage au pluriel qu'au singulier et qu'elle peut même :

[...] servir à les distinguer l'un de l'autre : il y a tendance (pas obligation) à marquer le pluriel par une liaison en [z] et le singulier par l'absence de liaison. (Delattre, 1947 : 149)

Il prend comme exemple, les paires suivantes (1947 : 150) :

- (8) *Des soldats_espagnols ; un soldat / espagnol*
- (9) *Des maisons_à vendre ; une maison / à vendre*
- (10) *Les lilas_et les roses ; le lilas / et la rose.*

Kaye et Morin (1982) confirment l'hypothèse de Gougenheim (1938) en testant le comportement de /z/ dans les trois phrases suivantes :

- (11) *Un marchand de draps_anglais*
- (12) *Un marchand de draps / anglais*
- (13) *Des marchands de draps_anglais*

Ils indiquent que dans des contextes de type Nom + Adjectif, /z/ fonctionne comme une motivation morphologique et pourrait être réanalysé comme un marqueur stylistique optionnel du pluriel. Cette hypothèse est renforcée par l'analyse des velours suivants :

- (14) *Des avions à réaction z-américains*
- (15) *Des chefs d'Etat z-africains*
- (16) *Les chemins de fer z-anglais*

Ce rôle grammatical se limite toutefois à cet unique contexte puisque, comme ils le font remarquer en reprenant les conclusions de Gougenheim :

[...] there is no liaison *z* before the complements of nouns as in *les participants / au congrès* [...]. It is precisely because these complements are not adjective-like and as such should not agree in number with the lexical head. (Morin et Kaye, 1982: 322)

Par ailleurs, Ågren (1963 : 124) signale que ce rôle grammatical a des limites, du fait même qu'il n'est possible que dans un contexte de liaison variable où, par définition, cette liaison n'est pas toujours réalisée :

Ce rôle fonctionnel qu'il remplit dans la langue écrite semble toutefois très précaire dans la langue parlée. Déjà, le fait même que la liaison du nom pluriel est facultative nous renseigne sur la valeur diminuée de cette fonction. Delattre préfère donner à ce 's' une fonction morphologique rendant possible (mais pas indispensable) l'opposition du pluriel et du singulier.

3.2.2 Les analyses par suite linéaire de catégories

Ce sont les analyses descriptives sur la liaison qui, les premières, ont établi l'importance de la syntaxe dans la réalisation de la liaison, ou plus précisément, dont on a pu dire qu'elles avaient établi l'importance de la syntaxe dans la réalisation de la catégorie.

La plupart de ces travaux fondent leur définition sur une analyse linéaire et catégorielle des éléments à lier. Et, dans le même temps, seuls deux éléments (M1 et M2) sont mentionnés : c'est dans la contiguïté que se joue le phénomène. Aussi, la combinaison syntagmatique complète de la phrase ou d'un groupe n'est-elle pas envisagée. La dimension syntaxique est, de ce fait, extrêmement réduite.

Le travail consiste alors à dresser des inventaires de combinaisons à deux éléments (catégories). Par exemple, « Déterminant + Nom » ; « Verbe + », *etc.*

Ainsi, Delattre fournit une liste détaillée de l'ensemble des combinaisons possibles. Par exemple :

1. Pour les liaisons catégoriques :

- Déterminatif + Nom (par ex. *tes amis*)
- Verbe + Pronom personnel (par ex. *Que veut-il ?*) ;

2. Pour les liaisons variables :

- Nom pluriel + Adjectif (par ex. *des maisons affreuses*)
- Adverbe polysyllabique + (par ex. *assez aimable*)

On remarquera en passant que les contextes identifiés par Delattre ne sont pas identifiés d'un point de vue formel comme purement syntaxiques : des éléments lexicaux peuvent être mentionnés, ou bien encore leur longueur. En outre, les catégories utilisées sont celles de la grammaire traditionnelle : les adjectifs possessifs sont classés avec les adjectifs qualificatifs, mais pas avec les déterminants, au titre qu'on les appelle tous deux « adjectifs ». Nous reviendrons en détail sur l'inventaire de Delattre dans la section 3.2.3.

On comprend alors qu'avec une syntaxe définie avec si peu de moyens (un jeu de deux catégories), on ait cherché à caractériser le lien entre les deux catégories mentionnées. C'est la mise en avant de la notion de cohésion syntaxique de la combinaison, soit le lien plus ou moins fort entre les deux mots. La cohésion syntaxique est posée comme déterminante pour la réalisation de la liaison : sans cohésion syntaxique, la liaison ne se fera pas. S'il y a cohésion, alors la liaison peut se faire.

C'est ce que Delattre (1955 : 45), à l'instar de Gougenheim (1935), appelle « *le degré d'étroitesse dans l'union des mots* » :

C'est l'union syntaxique très forte entre certains mots qui a maintenu « l'enchaînement ancien ». [...] Premier principe et seul nécessaire : la liaison se fait dans la mesure où l'usage a consacré l'extrême étroitesse d'union de deux mots ou classes de mots. (Delattre, 1966 : 40)

Cette notion de cohésion est déjà présente chez Nyrop (1934 : 128) qui indique que, dans la conversation, la cohésion est restreinte aux groupes de mots étroitement unis par le sens ou par l'accent. On dira *un profond accent* mais *cet ouvrage est profond / et juste*. Finalement, la liaison est réalisée « *surtout après les proclitiques, rarement après les mots accentués* ».

Le travail de caractérisation des contextes en termes de cohésion conduit à un résultat important, celui de la caractérisation des contextes où la liaison est obligatoire et ceux

où la liaison est variable. Certes, les indications, qui fourmillent dans ces travaux sont éparpillées et peu formalisées⁴⁴ : par exemple, Grammont (1914 : 130) fait observer que les mots « *accessoires* » se lient toujours quand ils sont préposés mais pas lorsqu'ils sont postposés. Selon Delattre, la liaison s'inscrit dans un rapport de force entre les deux termes à lier et se fait « *surtout en passant d'un mot moins important à un mot plus important* » (Delattre, 1947 : 149).

Delattre (1955) établit même un classement des contextes en fonction de leur degré de cohésion et les classe sur une échelle allant de 10 à 1, soit de la cohésion la plus forte à la plus faible. On trouve ainsi classés :

- en 10, les « Déterminatif + Nom », les « Adjectif + Nom » et les « Verbe + clitique », par exemple ;
- en 5, les « Nom + Adjectif », par exemple ;
- et en 1, les « Conjonction + », par exemple.

Toutes ces indications et ces analyses présentent l'avantage de pouvoir :

1. comptabiliser et classer les différents contextes de liaison dans les trois classes prévues, finalement déterminées de la manière suivante :

- les contextes où la liaison peut se réaliser (catégorique et variable) et ceux où elle ne peut pas avoir lieu (une partie des liaisons erratiques) ;
- puis, dans un deuxième temps, les contextes où la liaison est catégorique (cohésion forte) et ceux où elle est variable (cohésion moins forte).

On aboutit alors à la synthèse de Delattre (1966) sur laquelle nous reviendrons en détail dans la partie 3.2.3.2 ;

2. identifier les zones où la syntaxe spécifie de manière déterminante le caractère catégorique de certaines liaisons : il s'agit essentiellement de la zone à gauche de la tête syntaxique (pour employer une terminologie qui n'existait pas à l'époque) ; plus précisément, sera notée la partie à gauche du nom ou bien de l'entourage immédiat du verbe (avec les clitiques) ; certes les propositions cherchent à généraliser au-delà de ces

⁴⁴ Ågren est l'un des rares, à notre connaissance, à avoir cherché à fonder cette notion de manière fonctionnelle. Il reprend tout d'abord à son compte la notion de cohésion : « *Sans cohésion entre deux mots, il n'y aura pas de liaison. Donc, plus l'union des mots est forte, plus les chances sont grandes pour que la liaison se fasse.* » (Ågren 1973: 15). Puis il ajoute que la cohésion syntaxique est en partie délimitée par la fonction des deux termes : « [...] *il ne serait pas déraisonnable de supposer que, si la fonction du mot liant et/ou du mot lié ne change pas, la cohésion ne varie pas non plus* » (Ågren, 1973, p. 15).

deux grandes zones, mais comme nous le verrons par la suite, il ne semble pas que la dimension syntaxique permette d'aller plus loin.

Ces propositions n'ont, par contre, aucune valeur explicative ; elles ne permettent pas non plus d'étudier conjointement les différentes dimensions mises en jeu dans la liaison, ni de déterminer quelle position occupe la dimension syntaxique dans une hiérarchisation de contraintes, si tant est qu'il existerait une hiérarchie entre les différentes dimensions.

En outre, la priorité donnée à la partition des types de liaison relègue la description syntaxique du phénomène au second plan. Il s'agit finalement de caractériser les trois types de liaison. La dimension syntaxique en est le moyen. Mais elle n'est jamais envisagée pour elle-même. Dans les faits, la caractérisation syntaxique demeure seconde et peut s'exprimer dans une économie minimale : la consécution de deux catégories.

L'analyse de la dimension morphosyntaxique de la liaison travaillée dans la littérature a permis de spécifier (i) son rôle fonctionnel comme marqueur du pluriel ; (ii) de fonder la tripartition des contextes de réalisation ou de non-réalisation de la liaison. C'est avec une économie de moyens que ces contextes sont spécifiés : deux catégories consécutives.

Cette tripartition, et son fondement syntaxique, constitue un résultat important. En effet, cette tripartition n'a jamais été remise en cause. Par contre, son contenu a pu varier. C'est ce que nous allons examiner dans la section suivante.

3.2.3 Partition des liaisons

Les liaisons sont toujours définies selon des contextes, essentiellement de nature syntaxique comme nous venons de le voir. Leur recensement a été l'objet de beaucoup de travaux, notamment à partir des analyses de Delattre (1947 et 1955). Les résultats ont été le plus souvent présentés sous la forme de listes et, si le contenu de ces listes a évolué, la forme de la liste s'est, elle, maintenue.

Comme nous l'avons rappelé dans le Chapitre 2, nombreux sont ceux qui ont tenté de réguler le phénomène en imposant un modèle normatif. Dans la position classique et

normative (par ex. Fouché, 1959 ; Grevisse, 1969), il s'agissait d'imposer un schéma de référence. De cette volonté a émergé un système dans lequel il n'est plus question, en grande partie, que d'obligations et d'interdictions.

Dans cette section, nous reviendrons sur le contenu des contextes proposés. Nous utiliserons comme point de départ la grille tripartite établie par Delattre (1966). Premièrement, parce que cette grille constitue, aujourd'hui encore, la classification la plus détaillée ; deuxièmement, parce que son contenu a été modulé ou validé grâce à de nombreuses études contemporaines qui ont soumis le classement proposé à des analyses empiriques poussées (par ex. Ågren, 1973 ; Encrevé, 1988 ; De Jong, 1994, Morin, 1987 ; *etc.*).

Toutefois, avant de décrire le contenu et la motivation des contextes de liaison, il nous semble nécessaire de faire un point terminologique.

3.2.3.1 Terminologie

Tous les auteurs s'accordent pour dire qu'il existe trois types de liaison :

1. celles qui se réalisent toujours quel que soit le contexte ou le locuteur ;
2. celles qui ne se réalisent jamais quel que soit le contexte ou le locuteur ;
3. et celles qui se réalisent de façon plus ou moins variable selon le contexte et / ou le locuteur.

Concernant ces trois types de liaison, il existe deux ensembles de dénomination.

Tandis que certains parlent de « liaisons obligatoires », d'autres parlent de « liaisons catégoriques » (1.). De même, « liaisons interdites » alterne avec « liaisons erratiques » (2.) et « liaisons facultatives » avec « liaisons variables » (3.). Comme nous l'avons vu précédemment :

- (i) l'ensemble obligatoire / interdit / facultatif est issu des perspectives normativistes ;
- (ii) et l'ensemble catégorique / erratique / variable s'inscrit dans une perspective variationniste (*cf.* Labov, 1972 ; Encrevé, 1988 : 46).

Cette différence de terminologie n'est pas anodine. Elle illustre, à notre sens, les deux points de vue que nous avons décrits dans le Chapitre 2, à savoir une opposition

entre deux visions de la langue : (i) la volonté d'imposer des schémas de référence stricts, et (ii), la volonté de trouver une motivation aux usages de la langue.

Si cette tripartition, œuvre des normativistes, fait aujourd'hui l'objet d'un consensus, c'est parce que, comme le rappelle Lodge (1997), la codification mise en place par la norme s'appuie en partie sur le système « *sous-jacent* » de la langue, sur les structures linguistiques communes aux locuteurs d'une même communauté linguistique. Et cela, même si la finalité ne consiste qu'à :

- (i) décider quelles formes seront considérées comme « optimales » du point de vue de la « bonne société » et quels usages seront prohibés ;
- (ii) mais aussi à conditionner les instructions dictées par un « *idéal esthétique ou socioculturel d'un milieu détenant prestige et autorité* » (Garmadi, 1981 : 65).

Comme le souligne Lodge (1997 : 206) :

Codifier signifie donc produire des listes de règles. Mais une distinction importante doit ici être faite entre deux types de règles linguistiques : d'une part les règles *descriptives*, c'est-à-dire celles qui rendent explicites les structures sous-jacentes, le plus souvent subconscientes, de la langue (ce qu'en français on appelle souvent les lois du langage), et d'autre part les règles *prescriptives*, c'est-à-dire celles qui décident laquelle de deux formes linguistiques (ou plus) est à un moment donné considérée par la société comme correcte et acceptable (ce qu'en français on appelle souvent les règles). La codification fait intervenir des règles du second type, même si les grammairiens prescriptivistes s'efforcent généralement de présenter leurs règles comme s'il s'agissait de lois.

Pour Encrevé (1988), le problème posé par la tripartition des normativistes est avant tout d'ordre terminologique. Il ne remet pas en cause sa forme ou son contenu :

Délaissant le normatif pour le descriptif, il faudrait en toute rigueur rebaptiser selon le lexique de Labov les liaisons *obligatoires*, *facultatives* et *interdites* en liaisons *invariables* (ou *catégoriques*), *variables* et *erratiques*. [...] nous utiliserons les deux vocabulaires, mais, [...], c'est toujours le second qu'il faut entendre dans le premier : ce que nous visons, c'est la réalité de la pratique actuelle, et non les prescriptions qui peuvent être faites dans les traités. (Encrevé, 1988 : 46)

Aussi le travail d'Encrevé présente pour une part les caractéristiques d'une « entreprise de renomination » des classes de la liaison : la classification des

normativistes et la classification d'Encrevé restent, sur le fond et sur la forme, quasi-identiques.

Toutefois, la terminologie proposée par Encrevé est intéressante parce qu'elle fournit des dénominations cohérentes sémantiquement avec un cadre théorique variationniste. Elle induit un programme de questions qu'il était impossible de poser précédemment ou qui n'étaient pas au centre des travaux antérieurs.

Pour cette raison, nous adopterons la position d'Encrevé, soit l'utilisation de la terminologie variationniste, pour dénommer les trois classes de liaison. Pour simplifier l'exposé des sections qui suivent, nous utiliserons aussi la terminologie variationniste, y compris quand les auteurs ont eu recours à la terminologie normative.

3.2.3.2 Le contenu des classes de liaison selon Delattre (1966) et leur révision

Les propositions finales de Delattre sont résumées dans un ouvrage collectif de 1966 (*cf.* note de bas de page 9, p. 32). Le tableau qu'il propose regroupe les classes ou les sous-classes selon plusieurs paramètres :

- le caractère catégorique, variable ou erratique, qui vaut pour toutes les sous-classes ;
- la catégorie lexicale, même si Delattre n'emploie pas ce terme. On trouve le nom, le verbe, la préposition, l'adverbe et la conjonction⁴⁵ ;
- la flexion : on trouve à la fois singulier, pluriel et invariable ;
- une liste spécifique de mots ou d'expressions⁴⁶ ;
- un procédé de construction morphologique ;
- le nombre de syllabes : monosyllabique ou polysyllabique ;
- la position dans la séquence : suivi de ou précédé de (indiqué par un élément suivi ou précédé de +).

⁴⁵ Ces trois dernières catégories sont regroupées sous l'appellation « invariables » par Delattre.

⁴⁶ Ils se trouvent rassemblés sous l'appellation « spéciales » ou « invariables ».

La cohérence terminologique et les paramètres utilisés relèvent donc de plusieurs ordres d'organisation. Si le mélange de ces différents ordres est problématique, ce point ne fait pas question pour Delattre. Il apparaît plutôt que son objectif est l'exhaustivité.

	Obligatoires	Facultatives	Interdites
NOM	Déterminatif + - Nom <i>vos enfants</i> - Pronom <i>deux autres</i> - Adjectif <i>un ancien ami</i>	Nom pluriel + <i>des soldats anglais</i> <i>ses plans ont réussi</i>	Nom singulier + <i>un soldat anglais</i> <i>son plan a réussi</i>
VERBE	Pronom personnel + verbe <i>ils ont compris</i> <i>nous en avons</i> Verbe + pronom personnel <i>ont-ils compris</i> <i>allons-y</i>	Verbe + <i>je vais essayer</i> <i>j'avais entendu dire</i> <i>vous êtes invité</i> <i>il commençait à lire</i>	
INVARIABLES	Invariables monosyllabiques+ ⁴⁷ <i>en une journée</i> <i>trop aimable</i>	Invariables polysyllabiques+ <i>pendant un jour</i> <i>toujours utile</i>	<i>et +</i> <i>et on l'a fait</i>
SPECIALES	Formes figées <i>comment allez-vous</i> <i>les Etats-Unis</i> <i>accent aigu</i> <i>tout à coup</i> <i>de temps en temps</i>		'h' aspiré <i>des héros</i> <i>en haut</i> + <i>un, huit, onze</i> et dérivés <i>la cent huitième</i> <i>en onze jours</i>

Figure 3-6 Tableau simplifié de la classification des liaisons (Delattre, 1947 : 152)

Dans les sections suivantes, nous traiterons des évolutions qui ont été proposées pour les trois classes de liaison dans l'ordre suivant : les erratiques, les catégoriques, pour finir par les variables.

Nous commencerons d'abord par fournir le contenu détaillé de la classe étudiée, à partir des propositions de Delattre (1966). Nous n'utiliserons plus le symbole + mais expliciterons les listes en mentionnant la « catégorie » des termes M1 et M2 impliqués dans la liaison⁴⁸. Cela permettra de mieux distinguer quel est le terme organisateur de la liaison ou de la non-liaison et quelle est la position qu'il occupe.

⁴⁷ Il y a contradiction entre le tableau simplifié et le tableau détaillé proposé par Delattre (1947) puisque les « Conjonction monosyllabique + » sont catégoriques dans le tableau simplifié en tant qu'elles forment un sous-ensemble des « Invariables monosyllabiques + » et deviennent variables dans le tableau détaillé (cf. Figure 3-11, p. 76 et Delattre, 1947 : 155).

⁴⁸ Afin de simplifier la présentation, nous reprendrons les catégories utilisées par Delattre, même si certaines sont problématiques, notamment la classe des adjectifs qui est en fait très hétérogène ; au vu des travaux plus récents, elle regroupe aussi bien des éléments occupant une position Spécifieur (plus ou moins les déterminants) que des adjectifs qualificatifs.

Ensuite, nous ferons état des révisions qui ont été éventuellement suggérées, notamment grâce aux analyses quantitatives plus récentes. Afin d'être le plus clair possible, nous procéderons de la façon suivante : les révisions ne porteront tout d'abord que sur les contextes qui ont été éliminés de la classe traitée et non sur les nouveaux contextes à intégrer. Par exemple, si un contexte considéré comme erratique par Delattre a été par la suite considéré comme variable, alors il sera fait état de ce contexte dans la section « classe des erratiques », mais pas dans la section « classes des variables ».

Notre objectif est d'obtenir une classification des contextes de liaison qui tienne compte de l'ensemble des travaux réalisés. Aussi, c'est dans la section « Bilan » que nous rassemblerons l'ensemble des changements et que nous présenterons les nouveaux contenus des trois classes. Et c'est cette classification, objectivée par des listes de contextes, que nous confronterons aux données fournies par le corpus PFC (*cf.* 6.5, pp. 13 et *sq.*).

3.2.3.2.1 La classe des erratiques

Le tableau ci-dessous (Figure 3-7) rassemble l'ensemble des contextes de liaisons erratiques fournis par Delattre.

Lorsqu'une case est vide, cela signifie que n'importe quel terme peut apparaître. Ce procédé sera également utilisé pour la description des deux autres contextes de liaison (catégoriques et variables).

M1	M2	Exemple
nom singulier		<i>un chien errant</i>
nom propre		<i>François est au travail</i>
pronom (pas personnel) nasal		<i>le mien est tordu</i>
pronom personnel postposé nasal 3 ^{ème} personne	- pronom personnel sujets (<i>on, ils, elles</i>) - pronom complément (<i>en, les</i>)	<i>va-t-on en finir</i> <i>amusons-les un peu</i>
adverbe interrogatif		<i>quand es-tu tombé</i>
conjonction polysyllabique		<i>alors il est parti</i>
<i>et</i>		<i>et il l'a tué</i>
	<i>h aspiré</i>	<i>des hiboux</i>
	<i>oui, non</i>	<i>c'est oui</i>
- adjectif singulier - pronom nasal	invariable	<i>bon ou mauvais</i> <i>chacun à son tour</i>
	- <i>un, huit, onze</i> - les dérivés : <i>unième, huitième, onzième</i>	<i>dans les huit jours</i>
	nom composé au pluriel	<i>des moulins à vent</i>
	groupe figé sans liaison	<i>pot à tabac</i>

Figure 3-7 Tableau de la classe des liaisons erratiques

Aucun des contextes erratiques n'a jamais été remis en question.

En revanche, comme nous le verrons par la suite (« classe des variables »), il semble que l'un des contextes de liaison variable doive désormais intégrer la classe des erratiques.

3.2.3.2.2 La classe des catégoriques

Le tableau ci-dessous (Figure 3-8) rassemble l'ensemble des contextes de liaisons catégoriques fournis par Delattre :

M1	M2	Exemple
article - défini - indéfini - contracté - partitif		<i>les avions</i> <i>un ami</i> <i>aux amis</i> <i>des épinards</i>
adjectif - démonstratif - possessif - numéral - interrogatif - indéfini - qualificatif	nom	<i>ces étudiants</i> <i>mes envies</i> <i>trois enfants</i> <i>quelles aberrations</i> <i>certain animaux</i> <i>un petit animal</i>
déterminatif	- pronom - adjectif	<i>plusieurs autres</i> <i>certain anciens élèves</i>
pronom personnel	- verbe - pronom + verbe	<i>ils ont</i> <i>vous y êtes</i>
verbe	pronom personnel ou non	<i>mangez-en</i>
pronom personnel postposé au verbe	pronom	<i>allons-nous-en</i>
verbe impersonnel - <i>c'est</i> - <i>il est</i>		<i>c'est arrivé</i> <i>il est impossible</i>
préposition monosyllabique		<i>dans un an</i>
adverbe monosyllabique		<i>moins élégant</i>
groupes figés		<i>petit à petit</i>

Figure 3-8 Tableau des contextes de liaisons catégoriques

Contrairement aux liaisons erratiques, les liaisons catégoriques ont été révisées par différents auteurs. Il semblerait que quatre des contextes listés par Delattre, intègrent désormais la classe des liaisons variables : les *verbes impersonnels suivis d'un M2* ; les *adjectifs suivis d'un nom* ; enfin les *adverbes monosyllabiques suivis d'un M2* et les *prépositions monosyllabiques suivies d'un M2*.

1. Verbe impersonnel suivi d'un M2

Il semblerait que les constructions de type *verbe impersonnel suivi d'un M2* soient aujourd'hui variables.

Une étude de Durand et Lyche (2008) dans le corpus PFC⁴⁹ indique que seules 30 % de ces liaisons sont réalisées.

2. Adjectif suivi d'un nom

Il semblerait que les adjectifs préposés au nom manifestent une liaison variable. Post (2000), notamment, montre qu'en situation de lecture, la liaison entre un adjectif pluriel et un nom est variable et qu'elle peut tomber à 61 % de réalisation pour certaines combinaisons. Elle souligne cependant que, dans l'ensemble, cette liaison est majoritairement réalisée (88 %)⁵⁰. Toutefois, elle ne précise pas de quels types d'adjectifs il s'agit.

Nous distinguerons quant à nous les déterminants des adjectifs qualificatifs. Il est manifestement faux de dire qu'en français, on ferait des liaisons variables dans *ces enfants, mes enfants, trois enfants, etc.* Aussi, à ce moment de la présentation, si l'on doit poser l'hypothèse qu'une partie des adjectifs induit des liaisons variables, cela concernera en premier lieu les adjectifs qualificatifs.

Pour appuyer notre hypothèse, Encrevé (1988 : 48-49) relève que les « *adjectifs antéposés présentant une CL, c'est-à-dire masculins, appellent une précision* ». Le corpus de Laks (1977) montre qu'à l'exception d'un ensemble restreint (*petit, gros et grand*), les adjectifs masculins antéposés au nom présentent une liaison variable. De l'analyse de son propre corpus, Encrevé conclut que la liste des adjectifs antéposés se limite à quatre cas de figure : (i) ceux qui se terminent par une consonne nasale (par ex. *prochain, ancien*) ; (ii) ceux qui se terminent par *r* (par ex. *premier, dernier*) ; (iii) ceux qui ont une consonne finale fixe (*admirable, cher, etc.*) ; et (iv) ceux qui adoptent la forme féminine avec consonne finale fixe dans l'antéposition (par ex. *nouvel, vieil*).

De la même façon, les résultats de Durand et Lyche (2008) indiquent que les constructions de type *adjectif + nom* manifestent une liaison variable. Les exemples qu'ils donnent proviennent essentiellement de la lecture du texte du corpus PFC ; ce texte contient deux contextes : *grand émoi* et *grand honneur*. Sur cent locuteurs, six

⁴⁹ Pour leur article, Durand et Lyche (2008) ont exploité dix points d'enquêtes du corpus PFC (cf. pp. 159 et sq.), soient les données d'un peu plus d'une centaine de locuteurs du français.

⁵⁰ Ces résultats peuvent étonner lorsque l'on tient compte du fait que (i) la liaison est plus fréquente dans les situations de lecture que dans des situations de conversation spontanée (Fougeron et al., 2001), mais aussi (ii) qu'elle est plus fréquente au pluriel qu'au singulier (Morin et Kaye, 1982 ; Bonami, Boyé et Tseng, 2005).

n'ont pas réalisé ces liaisons, quatre dans le premier contexte, et deux dans le second. Outre le fait que ces résultats montrent la variabilité existant dans les combinaisons de type *adjectif suivi d'un nom*, surtout si l'on inclut les situations de lecture dans le registre « soutenu » (*cf.* pp. 80 et *sq.*), ils illustrent le problème que pose l'analyse des adjectifs préposés à la catégorie principale dans un contexte de liaison. Premièrement, les adjectifs susceptibles d'être antéposés sont très peu nombreux ; et deuxièmement, leur nombre est encore plus restreint en contexte de liaison. Par exemple, *gros* est répertorié 139 fois dans le corpus de conversation (de PFC), mais seules 8 occurrences se trouvent dans un contexte de liaison. Par ailleurs, la liaison n'est réalisée que dans 4 de ces contextes.

Au final, il semblerait donc que les adjectifs qualificatifs suivis d'un nom présentent une liaison variable et non pas catégorique, à l'exception de *petit*.

3. Adverbe monosyllabique suivi d'un M2

Les *adverbes monosyllabiques suivis d'un M2* induiraient désormais une liaison variable.

Cette proposition est soutenue par les résultats de Malécot (1975), qui indique que dans son corpus, *très* et *plus* ne manifestent la liaison qu'à hauteur de 63 % (111 réalisations sur un total de 176 contextes) et que *pas* ne manifeste la liaison qu'à hauteur de 37 % (16 réalisations sur un total de 43 contextes) (*cf.* également Ågren, 1973).

Encrevé (1988 : 48) intègre quant à lui tous les *adverbes monosyllabiques suivis d'un M2* à la classe des liaisons variables :

La modification porte sur un seul point. Soit que la position de Delattre ait été descriptivement inexacte, soit que la situation ait évolué, le domaine des liaisons invariablement réalisées par les « locuteurs cultivés » s'est restreint : les mots invariables monosyllabiques ont rejoint les mots invariables polysyllabiques dans la catégorie des liaisons réalisées variablement.

Les résultats de ces prédécesseurs combinés à ses propres résultats le font tout de même conclure qu'il est nécessaire de préciser la catégorie des invariables monosyllabiques, les prépositions et les adverbes ne manifestant pas le même type de liaison ; et que « la catégorie des mots invariables monosyllabiques que Delattre

classait parmi les obligatoires (les prépositions et les adverbes) est en fait disloquée sous ce rapport » (Encrevé, 1988 : 64). En effet, l'analyse de la liaison dans les débats télévisés de 1974 et 1981 entre Giscard d'Estaing (VGE) et Mitterrand (FM), montre que si *très* est réalisé catégoriquement par les deux figures politiques, *plus* est réalisé variablement par VGE (66,6 % de réalisation en 1974 et 50 % en 1981). L'adverbe *pas* est quant à lui réalisé variablement par VGE et FM (cf. Encrevé, 1988 : 65, Tableau 5) :

	1974	1981
VGE	16,7 %	37 %
FM	63,2 %	91,3 %

Figure 3-9 Réalisation de la liaison dans la construction « adverbe *pas* suivi d'un M2 » dans les débats de 1974 et 1981 entre VGE et FM (Encrevé, 1988 : 64 - 65)

Il semblerait donc que les *adverbes suivis d'un M2* rejoignent la classe des liaisons variables.

4. Préposition monosyllabique suivie d'un M2

Pour les *prépositions monosyllabiques suivies d'un M2*, les résultats des analyses postérieures à Delattre peuvent être contradictoires.

L'analyse d'Ashby (1981) révèle que seules 49 % (25 réalisations sur un total de 51 contextes) des prépositions monosyllabiques liaisonnent avec le mot qui suit.

Ces résultats sont toutefois remis en cause par une analyse quantitativement plus importante : celle de Malécot (1975), qui indique que, si la liaison avec une préposition monosyllabique est variable, elle ne l'est qu'à hauteur de 8 % (soit 168 réalisations sur un total de 182 contextes).

Durand et Lyche (2008), quand ils reprennent cette question, montrent que la liaison des *prépositions monosyllabiques suivies d'un M2* a bien intégré la classe des variables, à l'exception de la préposition *en*, qui continue de manifester une liaison catégorique :

Our own observations corroborate this statement and offer certain precisions: *en* induces a nearly categorical liaison, and out of 1124 relevant occurrences, we note two examples only without liaison: *en// un quart d'heure*, *manqué beaucoup de confiance en// elle*. All other prepositions vary in their usage, although liaison instances are overwhelming: there is no liaison after *dans* in 5 % of the occurrences,

including fixed expressions like *dans// un sens*. (Durand et Lyche, 2008 : 13)

Au vu des travaux mentionnés, nous pouvons raisonnablement penser que les *prépositions monosyllabiques suivies d'un M2* sont devenues des liaisons variables. La question pour la préposition *en* reste toutefois posée.

Si l'on tient compte de l'ensemble de ces résultats, le contenu de la classe des liaisons catégoriques se trouve modifiée. Les quatre points de discussion peuvent se résumer en deux cas de figure comme nous l'indiquons dans la Figure 3-10 :

1. les contextes qui passent de la classe des liaisons catégoriques à celle des liaisons variables ;
2. et les contextes qui se partagent entre les deux classes.

Contexte	Liaison catégorique	Liaison variable
adverbe monosyllabique suivi d'un M2		⇒
verbe impersonnel suivi d'un M2		⇒
préposition monosyllabique suivi d'un M2	<i>en</i>	toutes les autres prépositions monosyllabiques
adjectif suivi d'un M2	- déterminants - <i>petit</i>	- adj. qualificatifs pluriels - tous les autres adj. qualificatifs singuliers

Figure 3-10 Tableau des contextes « catégoriques » intégrant tout ou partie la classe des liaisons variables

3.2.3.2.3 La classe des variables

Cette classe regroupe tous les autres contextes que Delattre n'a pas intégrés dans les deux classes que nous avons vues précédemment.

Comme pour les deux autres classes, nous présentons ci-dessous le tableau qui rassemble l'ensemble des contextes de liaisons variables fournis par Delattre :

M1	M2	Exemple
nom pluriel	adjectif	<i>des avions énormes</i>
nom pluriel	verbe	<i>les parents attendent</i>
nom pluriel	invariable (préposition, adverbe)	<i>des projets à faire</i>
pronom non personnel		<i>toutes auront peur</i>
pronom personnel postposé au verbe	sauf pronom	<i>avons-nous soif bougez-vous un peu</i>
verbe		<i>je suis en transe</i>
	participe passé	<i>vous n'avez pas écouté</i>
	invariable	<i>bons ou mauvais</i>
préposition polysyllabique		<i>depuis un siècle</i>
adverbe polysyllabique		<i>souvent absent</i>
conjonction monosyllabique		<i>puis il est parti</i>

Figure 3-11 Tableau des contextes de liaisons variables

Un seul des contextes de liaisons variables proposés par Delattre a été remis en question.

Durand et Lyche (2008) excluent le contexte *noms pluriels suivis d'un verbe* de la classe des liaisons variables. Les résultats obtenus dans la tâche de lecture qu'ils ont soumise à leurs locuteurs annoncent que sur cent locuteurs, aucun n'a réalisé la liaison dans *Quelques fanatiques [z] auraient même entamé*. Leur proposition d'intégration de cette combinaison dans la classe des liaisons erratiques est aussi partagée par Bonami, Boyé et Tseng (2005 : 93), qui rappellent qu'il est « *généralement admis que la liaison est impossible entre un sujet syntagmatique et le verbe (les enfants ≠ arrivent)* ».

Il semblerait alors que les *noms pluriels suivis d'un verbe* rejoignent la classe des liaisons erratiques.

3.2.3.2.4 Bilan

Comme nous venons de le voir, les travaux ultérieurs à Delattre se sont essentiellement attachés à circonscrire l'espace de la variation, en notant que tel ou tel élément de telle sous-classe induit une liaison plus ou moins fréquente, et doit donc intégrer une autre classe que celle mentionnée.

Il en résulte que les trois classes de liaison ont subi des modifications importantes. Elles se sont enrichies et / ou appauvries de nouveaux contextes, et parfois les deux :

1. la classe des erratiques a intégré un nouveau contexte (*noms pluriels suivis d'un verbe*), mais aucune liaison erratique n'est devenue variable ou catégorique ;

2. la classe des catégoriques a été dépourvue de quatre contextes (*verbes impersonnels* suivis d'un M2 ; *adjectifs* suivis d'un *nom* ; *adverbes monosyllabiques* suivis d'un M2 et *prépositions monosyllabiques* suivis d'un M2). En revanche, aucun nouveau contexte n'est venu enrichir son contenu ;

3. enfin, la classe des variables a récupéré les quatre contextes « perdus » par la classe des catégoriques et a elle-même laissé un contexte, qui s'intègre désormais à la classe des erratiques (*noms pluriels suivis d'un verbe*).

En termes d'hypothèses à tester (*cf.* 6.5, pp. 271 et *sq.*), il s'agira pour nous de reprendre la nouvelle répartition des contextes de liaison, de la tester et d'éventuellement la modifier.

L'ensemble de ces révisions est reporté dans les tableaux ci-dessous (*cf.* Figure 3-12, Figure 3-13 et Figure 3-14) :

La nouvelle classe des erratiques :

M1	M2	Exemple
nom singulier		<i>un chien errant</i>
nom propre		<i>François est au travail</i>
pronom (pas personnel) nasal		<i>le mien est tordu</i>
pronom personnel postposé nasal 3ème personne	- pronom personnel sujets (<i>on, ils, elles</i>) - pronom complément (<i>en, les</i>)	<i>va-t-on en finir</i> <i>amusons-les un peu</i>
adverbe interrogatif		<i>quand es-tu tombé</i>
conjonction polysyllabique		<i>alors il est parti</i>
conjonction <i>et</i>		<i>et il l'a tué</i>
	h aspiré	<i>des hiboux</i>
	<i>oui, non</i>	<i>c'est oui</i>
- adjectif singulier - pronom nasal	invariable	<i>bon ou mauvais</i> <i>chacun à son tour</i>
	- <i>un, huit, onze</i> - les dérivés : <i>unième, huitième, onzième</i>	<i>dans les huit jours</i>
nom composé au pluriel		<i>des moulins à vent</i>
groupe figé sans liaison		<i>pot à tabac</i>
⇒ nom pluriel	Verbe	<i>les parents attendent</i>

Figure 3-12 Nouvelle répartition des contextes de liaisons erratiques

La nouvelle classe des catégoriques :

M1	M2	Exemple
article - défini - indéfini - contracté - partitif	nom	<i>les avions</i> <i>un ami</i> <i>aux amis</i> <i>des épinards</i>
adjectif - démonstratif - possessif - numéral - interrogatif - indéfini - qualificatif <i>petit</i>		<i>ces étudiants</i> <i>mes envies</i> <i>trois enfants</i> <i>quelles aberrations</i> <i>certain animaux</i> <i>un petit animal</i>
déterminatif	- pronom - adjectif	<i>plusieurs autres</i> <i>certain anciens élèves</i>
pronom personnel	- verbe - pronom + verbe	<i>ils ont</i> <i>vous y êtes</i>
verbe	pronom personnel ou non	<i>mangez-en</i>
pronom personnel postposé au verbe	pronom	<i>allons-nous en</i>
préposition <i>en</i>		<i>en un an</i>
groupes figés		<i>petit à petit</i>

Figure 3-13 Nouvelle répartition des contextes de liaisons catégoriques

La nouvelle classe des variables :

M1	M2	Exemple
nom pluriel	adjectif	<i>des avions énormes</i>
nom pluriel	invariable (préposition, adverbe)	<i>des projets à faire</i>
pronom non personnel		<i>toutes auront peur</i>
pronom personnel postposé au verbe	sauf pronom	<i>avons-nous soif bougez-vous un peu</i>
verbe		<i>je suis en transe</i>
	participe passé	<i>vous n'avez pas écouté</i>
	invariable	<i>bons ou mauvais</i>
préposition polysyllabique		<i>depuis un siècle</i>
adverbe polysyllabique		<i>souvent absent</i>
conjonction monosyllabique sauf <i>et</i>		<i>mais encore</i>
⇒ adverbe monosyllabique		<i>moins élégant</i>
⇒ verbe impersonnel		<i>c'est arrivé</i>
⇒ toutes les prépositions monosyllabiques sauf <i>en</i>		<i>dans un an</i>
⇒ adjectif qualificatif pluriel		<i>les nouveaux arrivants</i>
⇒ adjectif qualificatif singulier, sauf <i>petit</i>		<i>mon premier ami</i>

Figure 3-14 Nouvelle répartition des contextes de liaisons variables

3.3 Sociolinguistique

Comme nous l'avons indiqué en introduction de ce chapitre (*cf.* p. 37), l'analyse du phénomène de la liaison est complexe du fait que sa réalisation est à la fois sensible à deux ensembles de facteurs distincts : des facteurs linguistiques et des facteurs sociolinguistiques. L'objet de cette section est de faire le point sur les résultats des travaux qui se sont attachés à ce deuxième ensemble de facteurs.

Ces travaux s'intéressent d'abord et avant tout aux liaisons variables, soit des liaisons dont la réalisation n'est pas réglée par les seuls ordres phonologique ou syntaxique, par opposition aux liaisons catégoriques et erratiques. En effet, interviennent dans la réalisation / non-réalisation des liaisons variables des facteurs sociolinguistiques qu'il importe d'identifier.

Du fait de l'importance accordée à la variation, les travaux, dont nous allons faire état, revendiquent un cadre d'analyse variationniste. En d'autres termes, ils ont pris le contrepied des travaux qui ont considéré la liaison d'un point de vue normatif (*cf.* Chapitre 2, pp. 21 et *sq.*).

Nous nous attacherons tout d'abord à rappeler brièvement l'architecture variationnelle, pour ensuite décrire l'ensemble des facteurs sociolinguistiques susceptibles d'influer sur la réalisation de la liaison.

3.3.1 L'architecture variationnelle

Dès lors que l'on « *aborde les langues du point de vue de leur instabilité et de leur hétérogénéité interne* » (Gadet, 2001 : 3), que l'on cherche à expliquer et à formaliser des faits d'observation, des usages concrets de la langue, on se situe dans un cadre variationniste. Etudier un phénomène tel que la liaison nécessite particulièrement ce type de positionnement :

La liaison est un *indicateur social explicite*, un des rares lieux de la langue où les plus antivariationnistes des linguistes ont été amenés à reconnaître la variation sociale et l'hétérogénéité linguistique. (Encrevé, 1988 : 30)

Pour définir les différentes dimensions de la variation, nous reprenons ici la classification traditionnelle, telle qu'elle a été décrite par Coseriu (1969) et présentée par Gadet (2007). Elle est résumée dans le tableau suivant :

	Temps	Changement	Diachronie
Variation inter-locuteurs (selon l'usager)	espace	géographique, régional, spatial	diatopie
	société, communauté	social	diastatie
Variation intra-locuteurs (selon l'usage)	styles, niveaux, registres	situationnel, stylistique, fonctionnel	diaphasie
	canal	oral/écrit	diamésie

Figure 3-15 Représentation de la variation (Gadet, 2007 : 23)

Comme le montre ce tableau, la corrélation entre la structure sociale et la structure linguistique porte sur quatre dimensions qui sont réparties dans deux catégories distinctes : (i) la variation inter-locuteurs ; et (ii) la variation intra-locuteurs. Bien que cette représentation semble isoler les différentes dimensions les unes par rapport aux autres, il est nécessaire de préciser qu'il n'en est rien et que, bien au contraire, ces dimensions sont interdépendantes et « *interagissent en permanence* » :

Le découpage en types de variation laisserait entendre une discontinuité, alors que diatopique, diastatique et diaphasique interagissent en permanence [...] l'opposition [entre diastatique et diaphasique, *ndlr*] n'existe que dans la définition des termes : un même trait linguistique peut, en des progressions parallèles, correspondre à une position sociale favorisée, ou à un usage formel de distance⁵¹. (Gadet, 2007 : 24-25)

3.3.1.1 La variation inter-locuteurs

La variation inter-locuteurs, « *selon l'usager, soit différents individus selon des angles différents* » (Gadet, 2007 : 23) intègre trois dimensions :

1. la diachronie (*i.e.* variation temporelle ou historique) qui étudie les changements (plus ou moins importants) qui peuvent affecter une langue au cours des siècles et de son histoire ;

⁵¹ Gadet (2007 : 25) reprend comme exemple une phrase extraite d'un discours d'André Malraux : « *puissent les commémorations des deux guerres s'achever [r] aujourd'hui* ». Selon elle, la liaison avec *achever* pourrait aussi bien être réalisée par un locuteur fortement scolarisé que par un locuteur « *ordinaire* » qui se trouverait en situation formelle. Elle pourrait également être « *l'oralisation d'un discours* ».

2. la diatopie (*i.e.* variation géographique, régionale ou spatiale) qui étudie les usages d'une langue dans une région ou zone géographique plus restreinte ;

3. la diastratie (*i.e.* variation sociale et démographique) qui étudie les différences entre le parler de locuteurs appartenant à une même époque et une même zone géographique mais qui se différencient sur les plans démographiques et sociaux :

Etude de la diversité des façons de parler dans une communauté, rapportées à la diversité démographique ou sociale (Gadet, 2007 : 172)

Ainsi, les différences générationnelles (opposition jeunes / personnes âgées) peuvent être étudiées, mais aussi les différences sociolectales (selon la position sociale), les différences technoclectales (liées à une profession ou à une spécialisation), les différences entre les ruraux et les urbains, les différences selon le niveau d'études, *etc.*

3.3.1.2 La variation intra-locuteurs

La variation intra-locuteurs, « selon le répertoire d'un même locuteur dans différentes activités » (Gadet, 2007 : 23) intègre deux dimensions :

1. la diaphasie (*i.e.* variation stylistique ou situationnelle) qui étudie la diversité du répertoire d'un même locuteur dans différentes situations de communication :

Etude de la capacité des locuteurs à moduler leur façon de parler dans une communauté, rapportée à la diversité des localisations spatiales. (Gadet, 2007 : 173)

Il a ainsi été mis en évidence qu'une situation formelle (demande d'emploi, conférence devant un public, *etc.*) est caractérisée par l'emploi de formes normatives plutôt que familières. Inversement, une conversation plus informelle (d'ordre privé telle qu'une discussion entre amis) permet de se détacher des formes prescriptives. La diaphasie est traditionnellement représentée en termes de niveaux de langue. Ces niveaux de langue, appelés aussi registres ou styles, sont au nombre de quatre :

- (i) le registre soutenu (recherché, soigné, contrôlé)
- (ii) le registre standard (courant, commun)
- (iii) le registre familial (relâché, spontané)
- (iv) le registre populaire (argotique)

2. la diamésie (*i.e.* diversité de canal oral ou écrit) qui étudie les différences existant entre la production orale et la production écrite des usagers.

3.3.2 La variation sociolinguistique et la tripartition

L'étude de la liaison est intéressante d'un point de vue sociolinguistique du fait que la variabilité manifestée trouve en partie son origine dans l'interdépendance entre la situation de communication dans laquelle se trouve le locuteur et son profil socioculturel :

[...] les études de la liaison (en particulier celle dite « facultative ») laissent apparaître une grande difficulté (pour ne pas dire parfois impossibilité) à établir des règles strictes et incontestées, éloignées autant que possible de la subjectivité. Une constante fluctuation, décelable en synchronie, a fait de l'étude de ce phénomène, un problème « incontestablement [...] des plus épineux de la prononciation française » (Ågren 1973 : 1) et a abouti à rebuter de nombreux chercheurs. Cette variabilité, cette fluctuation, [...], peuvent être révélées à travers des productions situationnelles différentes pour un même sujet, ou encore selon la catégorie socioculturelle du locuteur. Il existe donc un aspect incontestablement dynamique dans le phénomène de liaison, [...] et qui dépasse le simple marquage d'un niveau stylistique [...]. (Lucci, 1983 : 231-232)

Cependant, cette variation ne fait pas appel à l'ensemble des dimensions présentées précédemment et elle concerne uniquement les liaisons variables⁵². Nous commencerons par « éliminer » les cas où l'on n'observe pas de variation.

3.3.2.1 Absence de variation : catégoriques et erratiques

Il est reconnu par l'ensemble de la communauté linguistique que les liaisons catégoriques ne présentent pas de variation puisqu'elles sont, en principe, invariablement réalisées par tous les locuteurs. Encrevé (1988), pour ne citer que lui, indique que :

⁵² L'utilisation de l'expression « liaison variable » quand il s'agit de parler de variation peut conduire à des formulations que l'on pourrait juger pléonastiques. Nous en sommes consciente, mais nous avons préféré conserver la terminologie (non normative) la plus utilisée. Il importe donc, en particulier dans les sections qui font état des travaux de sociolinguistique, de considérer cette expression comme une dénomination et non comme une description.

Les liaisons « obligatoires » sont, de l'avis général, invariablement réalisées par tous les locuteurs. Lorsque Henri Bauche, qui donne de la « langue populaire parisienne » une image délibérément pittoresque, énumère les liaisons qu'on y entend, il donne, cinquante ans plus tôt, une liste largement conforme à celle des liaisons que réalisent catégoriquement les informateurs de Malécot appartenant [...] à la « haute bourgeoisie parisienne ». (Encrevé, 1988 : 48)

D'un point de vue sociolinguistique, ces (non)-liaisons se caractérisent par le fait que :

- (i) elles ne sont pas socialement représentées comme des liaisons puisqu'elles ne créent « aucune distinction, ni positive, ni négative » (Encrevé, 1988 : 49) ;
- (ii) elles ne peuvent pas représenter des lieux de variation « pas plus sociale que stylistique » (Encrevé, 1988 : 49).

Cette absence de variation trouve son pendant dans le fonctionnement des liaisons erratiques qui elles, ne sont, en principe, jamais réalisées. De ce fait, nous intégrons la catégorie des liaisons catégoriques et la catégorie des liaisons erratiques dans un même ensemble, celui des liaisons invariables.

S'il arrive qu'un locuteur omette une liaison catégorique ou produise une liaison erratique, cette (non)-production sera considérée comme agrammaticale, dans le sens où elle ne relève pas du « possible de langue » (Milner, 1989).

3.3.2.2 Les variables comme lieu d'observation de la variation

A contrario, la variation sociolinguistique est observable dans le fonctionnement des liaisons variables. Cependant, cette variation ne se manifeste pas dans tous les contextes de communication et serait plus particulièrement le fait des locuteurs les plus scolarisés :

Plus le discours est familier, moins les liaisons facultatives sont réalisées. Les liaisons sont donc sociolinguistiquement d'autant plus signifiantes qu'elles sont plus rares. On peut dire, en suivant Encrevé, que la liaison est un « phénomène sociolinguistique inversé », car ce sont les locuteurs les plus scolarisés qui présentent sur elle le plus large système de variation. Aussi la constitution d'un corpus sur la liaison incite-t-elle à faire appel à des locuteurs favorisés, en situation surveillée, puisque, sur un corpus ordinaire, il n'y aura guère plus à signaler que le respect prévisible des liaisons obligatoires. (Gadet, 1989 : 52-53)

A mesure que le discours devient plus familier, le taux de réalisation des liaisons variables diminue. Les niveaux de langue impliqués dans la liaison variable ont principalement été décrits par les normativistes et les pédagogues du début du XXe siècle. Ces derniers se sont appliqués à fournir un cadre précis des situations dans lesquelles les liaisons étaient plus ou moins réalisées.

Dans une visée prescriptive (*cf.* pp. 26-30), les niveaux de langue sont décrits de façon dichotomique. Fouché (1959), par exemple, remarque que la liaison est le seul phénomène pour lequel il est nécessaire de distinguer deux niveaux qui s'opposent : la conversation courante et le style soutenu. Pour chacun de ces styles, il distingue ce qui peut être admis ou permis, de ce qui est proscrit. Dans une visée plus descriptive, Delattre (1966) différencie « *au moins* » quatre styles : la conversation familière, la conversation soignée, la conférence et la récitation de vers. Il précise pour chacun de ces registres la « proportion » de liaisons réalisées (1947 : 149) :

- dans la conversation familière, on ne fait pas ou presque pas de liaisons facultatives ;
- dans la conversation soignée, on en fait une petite proportion ;
- dans la conférence, on en fait beaucoup ;
- et dans la récitation de vers, on les fait toutes ou presque.

Il faut toutefois noter que ces descriptions portent essentiellement sur le parler des classes les plus cultivées :

Le choix du style, exigé par les circonstances, peut faire varier considérablement la fréquence des liaisons. « *A mesure que le ton s'abaisse et devient plus familier, on lie de moins en moins,* » dit Pierre Fouché ([...] 1935, p.52). [...] Ce qui précède s'applique au langage de la classe cultivée. A mesure que l'on s'éloigne de cette classe, le nombre de liaisons diminuent [...]. (Delattre, 1955 : 44)

Les liaisons variables fluctuent non seulement en fonction du style employé et du locuteur qui s'exprime dans un contexte particulier, mais elles sont également variablement réalisées par un même locuteur. Autrement dit, leur instabilité est à la fois inter et intra-locuteurs. Qu'elles soient ou non réalisées, cela n'affecte aucunement leur statut : elles restent dans l'ordre du grammatical et ne sont jamais considérées comme des « erreurs de production » :

Les liaisons dites « facultatives » sont celles que l'on peut aussi bien faire qu'omettre sans pour cela dépasser les bornes du langage accepté. (Delattre, 1955 : 43)

C'est en ce sens qu'elles s'opposent aux liaisons catégoriques et aux liaisons erratiques.

On dispose alors de deux « espaces » distincts lorsqu'il s'agit de traiter de la variation sociolinguistique :

- (i) un premier espace, celui des liaisons invariables (liaisons catégoriques et erratiques), dans lequel la variation n'est pas permise. Si l'on produit une liaison erratique ou si l'on ne produit pas une liaison catégorique, nous sommes conduits à considérer ces productions comme agrammaticales ;
- (ii) et un second espace, celui des liaisons variables, dans lequel la variation a lieu, puisque toute production (réalisant ou non la liaison) peut être considérée comme grammaticale.

La description des facteurs déterminants pour la fréquence de réalisation de la liaison ne concerneront donc que « l'espace » des liaisons variables. Ce que nous allons décrire dans la partie suivante.

3.3.3 Les facteurs sociologiques de la variation de la liaison

Différentes études descriptives ont déterminé que les liaisons variables (désormais LV) sont socialement réparties et que leur fréquence augmente avec « *la hauteur de style* » (cf. Encrevé, 1988 : 49-50). Dès 1892, Passy écrit à ce propos :

L'emploi des liaisons varie considérablement selon le style et selon les personnes. Dans le langage littéraire on lie beaucoup plus que dans le style familier ; mais ce sont surtout les instituteurs, et encore plus les personnes peu instruites essayant de parler « très bien », qui introduisent des liaisons en masse. (Passy, 1892 : 118)

La diaphasie et la diastratie sont donc les deux dimensions les plus influentes sur la variation de la liaison. Les principaux facteurs entrant en jeu sont résumés dans le tableau ci-dessous :

Diaphasique	registre ou facteur stylistique
Diastratique	profil socioculturel (catégorie socioprofessionnelle et niveau d'études)
	âge et sexe

Figure 3-16 Dimensions et facteurs de variation de la liaison

3.3.3.1 Le facteur stylistique

Les auteurs sont unanimes quant au fait que la fréquence de la LV varie selon les différents niveaux d'usage de la langue. Martinon (1914 : 356) explique que, même si les liaisons variables sont très sûrement livrées au goût de chacun, elles n'en demeurent pas moins dépendantes des circonstances et qu'il « *est évident qu'on en fait plus en lisant qu'en parlant, parce qu'en lisant on recherche la correction du langage* ».

Le niveau stylistique est même considéré par de nombreux auteurs comme le facteur le plus déterminant dans la réalisation de la liaison variable (Delattre, 1947 et 1955 ; Fouché, 1959 ; Ågren, 1973 ; Anderson, 1983). Selon eux, la LV dépend en premier lieu du style employé et se fait d'autant moins que le style est familier. Autrement dit, plus le registre de langue est élevé, plus les liaisons variables ont tendance à être réalisées⁵³.

Si l'on reprend les quatre registres décrits de la classification traditionnelle (*cf.* pp. 83 et *sq.*), on observe alors la distribution suivante :

⁵³ En 1983, une étude de Lucci indique que l'influence des différents styles de parole sur la réalisation de la liaison met également en jeu des débits de parole différents. Dans une étude plus récente, Fougéron *et al.* (2001) tentent de déterminer la portée du débit de parole sur la réalisation de la liaison en utilisant deux types de registres : la lecture oralisée et la parole spontanée. Leurs résultats montrent que le débit de parole n'a pas d'effet significatif sur la fréquence de la liaison en situation de lecture. Le pourcentage de liaisons réalisées ne varie que très peu. En revanche, les taux de réalisation de la liaison sont plus importants en situation de lecture que dans la conversation.

réalisation de LV	+	↑	registre soutenu (recherché, soigné, contrôlé)
			registre standard (courant, commun)
			registre familial (relâché, spontané)
	-		registre populaire (argotique)

Si l'on reprend maintenant les quatre registres proposés par Delattre, on obtient alors la distribution suivante :

réalisation de LV	+	↑	récitation de vers
			registre de conférence
			registre soigné
	-		registre familial

Bien que l'on observe le même sens dans les distributions, les deux descriptions ne se situent pas exactement sur le même plan puisque :

(i) Delattre ne propose aucun registre « populaire ». Comme son corpus a été élaboré à partir du discours des classes cultivées, il n'y « voit » aucun registre populaire. Fouché (1959), quant à lui, ne propose que deux niveaux : un soutenu et un courant ;

(ii) si Delattre fait mention d'un registre « familial », le registre « standard » est quant à lui exclu. Selon la tradition grammaticale, le registre « familial » correspond à du spontané, du relâché. Le « standard » est décrit par Gadet (2007 : 28) comme étant en quelque sorte le registre « idéal » à atteindre par des locuteurs ayant un statut social élevé. Il a une fonction d'homogénéisation et jouit d'une réputation de prestige. Elle précise cependant que ce n'est en fait pas « *une variété parmi d'autres : ni usage effectif ni vernaculaire de qui que ce soit* » mais une « *construction linguistique et discursive, homogénéisante* ». Aussi peut-on faire l'hypothèse que le niveau « familial » distingué par Delattre correspondrait aux deux niveaux « familial » et « standard », et que le registre « soigné » correspondrait au niveau « soutenu » de la classification traditionnelle ;

(iii) Delattre fait état de deux registres qui ne sont pas prévus par la description traditionnelle du fait qu'ils correspondent plus à de l'écrit oralisé qu'à de l'oral : celui de la récitation de vers et celui de la conférence.

Aussi, si, pour le fonctionnement de la liaison, on cherche à établir une corrélation entre les registres décrits par Delattre, voire même ceux de Fouché et la représentation traditionnelle des registres de langue, on obtient alors le tableau suivant :

	Représentation traditionnelle des niveaux de langue	Fouché (1959)	Delattre (1966)
ORAL / ECRIT	/	/	récitation de vers
	/	/	conférence
ORAL	soutenu	soutenu	soigné
	standard	courant	familier
	familier		
	populaire	/	/

Figure 3-17 Identification des registres dans la représentation traditionnelle, chez Fouché (1959) et chez Delattre (1966)

Nous avons vu qu'il était entendu que, selon le style utilisé, la LV est plus ou moins réalisée. Toutefois, la possibilité d'utiliser tel ou tel style est dépendante du profil socioculturel du locuteur.

3.3.3.2 Le profil socioculturel

Les études qui se sont attachées à relier liaison et profil socioculturel n'ont pas, dans l'ensemble, précisé ce qu'elles entendaient par « socioculturel ». L'abondance des dénominations utilisées pour circonscrire les ensembles visés le montre particulièrement : par exemple, « bourgeoisie », « société cultivée », « classe populaire », « classe favorisée / défavorisée », « hommes politiques », *etc.* De plus, ces dénominations visent des réalités d'ordre différent et s'inscrivent dans des cadres d'analyse différents (classe sociale *versus* classe professionnelle ; cultivée *versus*

populaire, cultivée *versus* classe professionnelle, *etc.*). De manière générale, on ne sait pas précisément si ces dénominations sont objectivées par rapport à tout ou partie d'un ensemble professionnel, d'un niveau d'études, d'un niveau de vie, d'un contexte de vie, *etc.*

Face à ces difficultés, force est de constater que (i) les auteurs ont généralement adopté un fonctionnement « lâche » de catégories finalement floues, (ii) tout en se rabattant généralement sur des objectivations quantifiables, telles que le niveau d'études, la profession, *etc.* Ce qui pose tout autant problème, ainsi que le rappelle Gadet :

Les sociolinguistes ont tendu à privilégier le rôle d'indices sociaux quantifiables [...] Sur quels critères classer les locuteurs, ne serait-ce qu'en trois groupes, comme « classe ouvrière », « classe moyenne » et « classe supérieure » ? Trois facteurs sont souvent exploités comme indices de position sociale, justement ceux qui permettent une quantification : niveau d'études, profession [...] et type d'habitat [...]. Mais de tels classements, [...] vont-ils s'avérer pertinents pour tous les phénomènes (socio)linguistiques ? (Gadet, 2007 : 92)

Notre objectif n'est pas ici d'en discuter plus avant. Aussi mettons-nous de côté ces difficultés définitoires pour nous attacher aux résultats.

Le travail mené dans un cadre sociolinguistique a conduit à proposer que (i) les locuteurs de niveau socioculturel élevé ont dans leur usage la maîtrise de tous les registres, notamment donc celle du registre soutenu ; (ii) c'est chez ces locuteurs que la réalisation de la LV peut être observée. Or, c'est dans le registre soutenu qu'on observe le plus de liaisons variables réalisées (*cf.* le point précédent).

Ces deux points ont permis d'établir que premièrement, la liaison fonctionne comme différenciateur social. Elle est un « *indicateur sociolinguistique très fort* » (Gadet, 2001 : 51) qui permet de déterminer le profil social des locuteurs. Comme le rappelle Bourdieu (1979) de façon plus large, la langue permet d'évaluer assez sûrement la position sociale de celui qui en fait usage :

En fait, l'usage de la langue, c'est-à-dire aussi bien la manière que la matière du discours, dépend de la position sociale du locuteur qui commande l'accès qu'il peut avoir à la langue de l'institution, à la parole officielle, orthodoxe, légitime [...]. (Bourdieu, 1979 : 107)

Deuxièmement, la liaison est un « *phénomène sociolinguistique inversé* » (Encrevé, 1988 : 45), « inversé » en ce sens que le phénomène s'observe pour une catégorie sociale non attendue. En effet, on rapproche traditionnellement l'observation de la variation avec un défaut de maîtrise des codes, celui que l'on impute généralement aux classes sociales les « moins favorisées ». Ici, c'est bien l'inverse que l'on observe, puisque ce sont les « plus favorisés » qui réalisent le plus de liaisons variables.

3.3.3.3 L'âge et le sexe

Les études qui ont manipulé les données sociodémographiques ont pris en compte essentiellement l'âge et le sexe des locuteurs.

A notre connaissance, celles qui se sont attachées au sexe, sont au nombre de trois. Pour la première (Malécot, 1975 : 169), les femmes réaliseraient plus de liaisons que les hommes (67% contre 62 %). Pour la seconde (Ashby, 1981 : 50), c'est l'inverse : 36 % de liaisons réalisées chez les hommes, contre 32 % chez les femmes). Quant à la dernière étude (Ranson, 2008), elle montre qu'il n'y a pas de différence de réalisation entre les femmes et les hommes. Ces résultats, contradictoires (si tant est par ailleurs que les différences dans les résultats de chacune de ces études soient significatives), semblent donc indiquer que le sexe n'est pas un paramètre influent sur la réalisation de la liaison.

Pour le paramètre de l'âge, nous pouvons indiquer les résultats de trois études (Malécot, 1975 ; Ashby, 1981 ; Ranson, 2008), qui s'accordent chacune pour conclure que les locuteurs les plus âgés réalisent plus de liaisons variables que les locuteurs les plus jeunes.

Contrairement au sexe, l'âge semble donc être un paramètre influent sur la réalisation de la liaison.

Dans cette section, nous avons analysé l'espace des liaisons variables, pour lesquelles toute production est grammaticale (*cf.* pp. 83 et *sq.*). Nous avons laissé de côté l'espace des liaisons invariables (liaisons catégoriques et erratiques) : comme ces dernières peuvent donner lieu à des productions agrammaticales, elles n'intègrent pas le paradigme variationnel.

3.3.4 Les productions agrammaticales

De nombreuses « productions agrammaticales », le plus souvent appelées des « erreurs », sont réalisées par les locuteurs. Ces productions agrammaticales sont, rappelons-le, de deux types :

- la non-production d'une liaison catégorique ;
- et la production d'une liaison erratique.

3.3.4.1 Dans les contextes de liaison catégorique

Les liaisons catégoriques ne constituent pas, par définition, un lieu de variation et par suite, un lieu de distinction socioculturelle. Aussi lorsque, exceptionnellement, une liaison catégorique n'est pas produite, cette non-production n'est pas imputée à un manque de maîtrise des codes, mais simplement à une erreur occasionnelle et non reproductible. Si, par contre, on observe qu'une liaison catégorique est non réalisée et ce, de manière relativement régulière, son contexte tend alors à intégrer la classe des variables (*cf.* 3.2.3 et 3.2.3.2.2).

Aussi, la production des liaisons catégoriques a finalement été très peu étudiée. Elle n'a pas constitué l'un des thèmes importants de la recherche sur la liaison.

3.3.4.2 Dans les contextes de liaison erratique

La plupart des travaux qui se sont attachés aux liaisons agrammaticales portent principalement sur l'analyse des liaisons erratiques. Comme le rappelle Gadet (2007 : 65) « *l'intérêt sociolinguistique de la liaison se manifeste aussi dans les fausses liaisons, où est souvent en jeu l'hypercorrection* ». Si les liaisons variables sont plus souvent réalisées chez les locuteurs les plus scolarisés, les « fautes de liaison » se révèlent quant à elles beaucoup plus nombreuses dans le français populaire : « *Ce n'est donc pas par la présence ou l'absence de liaison que l'on pourra caractériser le français populaire, mais plutôt par les fautes* » (Gadet, 1992 : 47). La production de ces liaisons erratiques joue alors un rôle socialement différenciateur (Gadet, 1989 : 55).

Ces liaisons, que l'on appelle également « liaisons mal-t-à-propos » ou « fausses liaisons », se manifestent tout particulièrement dans des situations qui favorisent l'hypercorrection.

Ces « fausses liaisons » sont généralement appelées pataquès, cuirs et velours.

- **Le pataquès** est, selon la légende (Domergue, 1805 : 465), de formation imitative par ironie, d'après *ce n'est pas-t-à moi, je ne sais pas-t-à qui est-ce*. Il désigne une faute de langage qui consiste à substituer un *t* final à un *s* final ou inversement (*cuir*) et plus généralement une consonne qui n'existe pas à la finale du mot précédent (*velours*).

- **Le velours** est la production d'une liaison abusive avec une consonne de liaison (*z*) inexistante dans la forme de base du premier mot (par ex. *va-/z/-à-lui* ; cf. Mounin, 1974).

- **Le cuir** est également considéré comme une liaison abusive. Il consiste à remplacer la consonne de liaison par une autre lorsque la liaison est produite. Il consiste le plus généralement à entendre un *z* à la place d'un *t* (par ex. *tu n'étais point-/z/-ici*).

Selon Gadet (1989 : 55-56), le phénomène d'hypercorrection met en jeu à la fois l'ordre syntagmatique et l'ordre paradigmatique. Les pataquès syntagmatiques sont les produits « *d'une prégnance discursive interne à la chaîne, dont le mécanisme est plus ou moins complexe* » (par ex. *les auteurs* et *les non [z] auteurs*). Les pataquès paradigmatiques sont produits « *par prégnance de mémoire* » (par ex. *si vous laissez [r] un message* en analogie avec *laisser [r] un message*).

Une autre façon de caractériser ces productions agrammaticales est de considérer qu'il existe des « *erreurs morphologiques* », des « *erreurs lexicales* » et des « *erreurs à distance* » (cf. notamment Desrochers, 1994) :

1. Les erreurs morphologiques

Produire une erreur de « type morphologique », essentiellement flexionnel, consiste à insérer une consonne fautive. Cette consonne semble jouer un rôle de marqueur morphologique soit du pluriel (par ex. *vingt [z] articles*), soit de l'impératif (par ex. *donne-moi [z] en*), soit verbal (par ex. *Malbrough s'en va [t] en guerre*).

2. Les erreurs lexicales

Une erreur de « type lexical » correspond à l'insertion d'une consonne de liaison dans un mot qui ne contient pas de CL, par analogie avec un autre mot de la même catégorie. Cet autre mot apparaît quant à lui avec cette consonne dans un contexte identique (par ex. *malgré [z] une blessure au genou* et *sans[z] une blessure au genou*).

3. Les erreurs à distance

Enfin, dans une erreur « à distance », la consonne de liaison est séparée du mot auquel elle appartient par d'autres éléments lexicaux (par ex. *qui est désormais [t] interdit ; j'ai beaucoup pensé [p] à toi*).

Côté (2005 : 71) remarque que, pour « *tous les locuteurs canadiens consultés* », ces erreurs « *frappent par leur naturel* », et suggère, à la suite de Armstrong (2001), « *qu'elles [ces « erreurs »] ne sauraient correspondre à de simples erreurs de « performance* » » contrairement à ce que suggère notamment Gadet (1989). Le problème que posent ces productions agrammaticales est qu'il est compliqué de distinguer ce qui relève d'un usage non maîtrisé du fonctionnement de la liaison de ce qui relève d'une grammaire des locuteurs qui ne correspond pas au schéma imposé.

3.3.5 Bilan

Envisagée d'un point de vue sociolinguistique, la liaison est vue avant tout comme un phénomène variable. Cette variation est inter ou intra-locuteurs. Elle se manifeste dans l'espace des liaisons variables. Les liaisons catégoriques et erratiques ne sont pas considérées.

Il apparaît que les dimensions les plus influentes sur la réalisation de la liaison variable sont la diaphasie (facteur stylistique) et la diastratie (profil socioculturel et âge

du locuteur). Dans la diastratie, le profil socioculturel du locuteur joue le rôle le plus important avec toutefois une difficulté à l'objectiver (*cf.* Encrevé, 1988 : 258-262). Il ressort également que ces deux dimensions ne sont pas totalement indépendantes, la diastratie étant partie liée avec la diaphasie qui est, quant à elle, la dimension prépondérante sur toutes les autres.

Les principaux résultats indiquent que la liaison est un différenciateur social très fort et un *phénomène sociolinguistique inversé*. Ces résultats sont formulés à un niveau de généralité important et on peut s'interroger sur leur portée. Ils ne donnent en effet pas un accès direct à l'explication de la variation pour telle communauté ou tel locuteur ; ils se situent du côté de l'usage de la langue mais pas du côté de la prédiction de la réalisation de la liaison. Si tous les éléments sont repérés, ils ne sont ni articulés, ni pondérés et sont, le plus souvent, analysés successivement.

Dans l'espace des invariables (catégoriques et erratiques), on compte un certain nombre de productions agrammaticales : les liaisons catégoriques non réalisées et les liaisons erratiques réalisées. Nous avons vu que la sociolinguistique porte un intérêt tout particulier à l'analyse des « fausses liaisons » (erratiques réalisées) étant donné qu'un certain nombre de ces productions sont linguistiquement motivées (morphologie, lexique, *etc.*). Les liaisons catégoriques non réalisées sont, quant à elles, le plus souvent laissées pour compte. En effet, si un contexte de liaison catégorique manifeste une forte variation dans sa réalisation, alors ce contexte intègrera la classe des variables. Si, en revanche, les résultats n'indiquent aucune récurrence dans la non-réalisation d'un contexte catégorique, mais plutôt des faits isolés, alors ces productions seront vues comme de simples « erreurs » de la part du locuteur.

3.4 Lexicale

En sus des dimensions phonologique, morphosyntaxique et sociolinguistique, la liaison est soumise à un ensemble de contraintes relevant du lexique.

Nous avons partiellement abordé cette dimension par l'indication de la catégorie (en ce qu'elle était indicateur du contexte syntaxique).

Ici, nous regroupons sous lexical, les dimensions qui reposent sur (i) l'examen de la longueur de la forme du lexème, et (ii) de la fréquence de l'unité dans la langue.

3.4.1 La longueur

Dans la section 3.2, nous avons vu que la syntaxe permettait de départager les contextes : certains sont catégoriques (la liaison est toujours réalisée) et c'est la « cohésion » forte qui assure une liaison catégorique. Cependant l'approche syntaxique ne détermine qu'une partie des contextes. Il en reste un grand nombre pour lesquels on observe ou non une liaison réalisée. De plus, on peut avoir à l'intérieur d'une même catégorie syntaxique des comportements différents selon les éléments qui constituent la catégorie.

Les travaux se sont alors attachés à trouver et à décrire d'autres paramètres pouvant influencer la réalisation de la liaison. Ils se sont notamment intéressés à la longueur des éléments à lier. La réalisation de la liaison serait dépendante de la longueur des termes : plus ils sont courts, plus la liaison aura tendance à se réaliser.

La proposition, dont nous verrons ci-dessous qu'elle se décline en fait en 3 hypothèses distinctes, peut être illustrée de la manière suivante :

(17) *De beaux_enfants*

(18) *De magnifiques_esquisses*

En (17), l'adjectif *beaux* est monosyllabique alors qu'en (18), l'adjectif *magnifique* est polysyllabique. La probabilité de réalisation de la liaison est donc plus forte en (17) qu'en (18).

La ligne de ces travaux s'appuie en fait sur les classements originaux de Delattre, qui faisaient apparaître une différence entre les « invariables » monosyllabiques et les « invariables » polysyllabiques, les premiers conduisant à une liaison catégorique, les seconds à une liaison variable (cf. 3.2.3.2, pp. 67-76)⁵⁴. Le travail ultérieur, comme on l'a vu a largement consisté à tester les contextes de Delattre, notamment ces contextes

⁵⁴ La seule exception porte sur les conjonctions : les monosyllabiques intègrent la classe des liaisons variables, les polysyllabiques, celle des liaisons erratiques. On retrouve donc le même type de contrainte.

identifiés par la longueur des mots (*cf.* entre autre, Encrevé, 1988). Si la critique des tableaux de Delattre a été le point de départ, l'examen de l'influence de la longueur est devenu un paramètre à tester en tant que tel.

De manière plus générale, cette proposition, soit l'influence de la longueur sur la réalisation de la liaison, ne concerne, comme nous l'avons rappelé ci-dessus, qu'un certain nombre de cas que la syntaxe n'a pas réglé de façon catégorique. Aussi, l'hypothèse de l'influence de la longueur, ne vaut-elle que pour les contextes dans lesquels il reste à déterminer s'ils sont catégoriques ou variables. Elle ne touche pas, *a priori*, des contextes du type « Déterminant + N » (par ex. *les enfants, certains amis* sont considérés comme des liaisons catégoriques), pour lesquels la cohésion syntaxique est maximale.

Parmi les auteurs qui ont cherché à relier la longueur et la réalisation de la liaison, on peut distinguer trois propositions :

- (i) la réalisation de la liaison est fonction de la longueur de M1 ;
- (ii) la réalisation de la liaison est dépendante de la longueur du groupe syntagmatique auquel appartient M2 ;
- (iii) la réalisation de la liaison est dépendante de la longueur du groupe syntagmatique auquel appartient M1.

Nous examinons successivement ces 3 propositions.

3.4.1.1 La longueur de M1

L'hypothèse la plus souvent soutenue est la suivante : plus M1 est court, plus la probabilité de réalisation de la liaison augmente. En d'autres termes, la liaison est plus souvent réalisée après un monosyllabique qu'après un polysyllabique.

Delattre (1947) rend compte du fait que la fréquence de la liaison est le résultat direct d'un conditionnement multifactoriel. Selon lui, la syntaxe n'est pas le seul ordre d'organisation déterminant les contextes de liaisons catégoriques et de liaisons variables. La longueur du lieu (M1) est également déterminante : « *la liaison se fait d'autant plus que le premier des deux mots est plus court* » (1947 : 150). Il observe que

ceci est particulièrement vrai avec les « invariables » (prépositions et adverbes) qu'il répartit dans les classes de liaisons catégoriques et de liaisons variables selon leur longueur (cf. 3.2.3.2.2 et 3.2.3.2.3) : un invariable monosyllabique (par ex. *trop_anxieux*) intègre la classe des catégoriques et un invariable polysyllabique celle des variables (par ex. *tellement_anxieux*).

Les résultats d'une étude menée par Fougeron *et al.* (2001a) confirment que la fréquence de réalisation de la liaison est plus forte lorsque M1 est court. Également, les données d'Encrevé (1988) indiquent que, toutes catégories lexicales confondues, les liaisons après un monosyllabe sont réalisées dans 77,8 % des cas. Les liaisons après un polysyllabe ne sont réalisées que dans 29 % des cas. Et l'analyse des contextes syntaxiques montre un comportement extrêmement hétérogène de la liaison à l'intérieur d'une même catégorie (cf. pp. 72-75). S'interrogeant alors sur l'interprétation de ces résultats, il conclut que si la syntaxe permet de circonscrire l'ensemble des contextes de liaisons catégoriques et variables, elle ne constitue aucunement le facteur le plus déterminant dans la réalisation de la liaison variable :

[...] le paramètre le plus décisif pour la réalisation d'une liaison facultative, ce n'est pas la catégorie syntaxique du premier mot, mais sa nature prosodique⁵⁵, sans que cela remette en cause, évidemment, les conditions de liens syntaxiques nécessaires pour qu'une liaison soit *possible* entre deux mots. (Encrevé, 1988 : 67-68)

De la même façon, l'analyse des différentes formes du verbe *être* par De Jong (1994) montre que la longueur est un des facteurs exerçant une influence sur la réalisation de la liaison : les formes fléchies monosyllabiques réalisent la liaison dans 61,5 % des cas, alors que les formes fléchies polysyllabiques ne réalisent la liaison qu'à hauteur de 15 %.

L'hypothèse de la longueur de M1 comporte toutefois un certain nombre de faiblesses.

⁵⁵ Dans cet extrait de son ouvrage dans lequel il parle de « *nature prosodique* », Encrevé (1988 : 67) ne fait mention que de longueur de mots.

Premièrement, la « révision » du contenu de la grille tripartite de Delattre (cf. 3.2.3.2, pp. 67 et sq.) a montré que de nombreux « invariables » monosyllabiques tels que prépositions et adverbes, ne présentaient en fait qu'une liaison variable.

Deuxièmement, les résultats obtenus par De Jong sont tempérés par le fait que la variation est plus ou moins fréquente selon les unités en jeu. Outre le fait qu'il ne s'agit que de formes de *être*, De Jong met par exemple en évidence que la liaison avec *est* se révèle beaucoup plus fréquente qu'avec *sont*. Avant lui, Ågren (1973 : 44) observe que, dans ce même contexte syntaxique (*être* + *un* Nom), la fréquence de réalisation de la liaison de lexèmes de même longueur est plus ou moins forte : la liaison dans « *est* [t] + *un* Nom » est réalisée quasi-systématiquement (98,8 %), alors qu'elle est très variable dans « *suis* [z] + *un* Nom » (53,8 %).

Ces résultats affaiblissent donc la proposition selon laquelle la longueur de M1 est déterminante pour la réalisation de la liaison.

3.4.1.2 La longueur du groupe syntaxique auquel appartient M2

Concernant la classe des variables, Delattre (1955) affirme que la longueur de M1 n'est pas la seule à influencer la réalisation de la liaison : la longueur du groupe syntaxique auquel appartient M2 joue également un rôle dans sa fréquence de réalisation : la liaison est plus probable dans un contexte tel que *tes enfants_ont grandi* (3 syllabes) que dans *tes enfants_ont tellement grandi* (5 ou 6 syllabes).

Ainsi, la distinction ne porte pas seulement sur une opposition entre monosyllabes et polysyllabes mais sur le nombre plus ou moins important de syllabes du groupe syntaxique dont est membre M2. La liaison sera plus fréquente lorsque ce groupe sera bisyllabique (par ex. *j'avais_une vie*), que lorsqu'il sera trisyllabique (par ex. *j'avais_une voiture*), etc. :

Bisyllabique	Trisyllabique	Quadrisyllabique etc.
Liaison + fréquente		Liaison - fréquente

Morin et Kaye (1982 : 296) font eux aussi l'hypothèse que les liaisons semblent plus fréquentes avec un « complément court » (19) qu'avec « un complément long » (20) :

[...] we have shortened the length of the complement to be tested in some of examples, since it seemed to us, but it should be verified, that liaison easier before short complement than before long ones.

(19) *Ils travaillent d'abord et mangent après*

(20) *D'habitude, ils mangent après avoir fini leur partie d'échecs*

Dans l'exemple (19), le complément est constitué d'une seule unité lexicale qui est bisyllabique. En (20), le complément est constitué d'une suite d'unités, qui totalisent 11 syllabes. La probabilité de réalisation de la liaison serait plus forte en (19) qu'en (20).

Comme le disent Morin et Kaye (1982), cette hypothèse doit être testée (« *should be verified* »). Cela signifie que, dans des contextes tels que « Verbe + » (repris de Morin et Kaye, mais en faisant varier les lexèmes ou bien le type de lien entre le Verbe et ce qui suit le verbe) ou bien encore dans un contexte « Nom pluriel + », on devrait observer des différences dans les probabilités de réalisation de la liaison :

- « Verbe + » : en (21), le groupe syntaxique auquel appartient M2, est plus court qu'en (22), alors la probabilité de réalisation de la liaison serait plus forte :

(21) *Le garçon veut une pomme*

(22) *Le garçon veut une autre pomme*

- « Nom pluriel + » : la probabilité de réalisation de la liaison serait plus forte en (23) qu'en (24) :

(23) *Les gens impatients*

(24) *Les gens impatients et anxieux / de faire leurs courses*

Au vu des exemples (21)-(24), il ne nous semble pas immédiat que cette hypothèse soit validée. Dans tous les cas, on voit aussi la difficulté à la tester précisément.

A la suite de Morin et Kaye (1982), mais en interprétant leur proposition de manière plus restrictive, Fougeron *et al.* (2001b) n'ont pas testé la longueur du groupe

syntaxique auquel appartient M2 mais simplement la longueur de M2. Leurs résultats font apparaître qu'en dehors des contextes « + Adjectif » pour lesquels, effectivement, la longueur de l'adjectif semble être un paramètre déterminant, la longueur de M2 n'a que peu d'influence sur la réalisation de la liaison. Pour une même catégorie syntaxique contenant des lieux (M2) de longueurs différentes, les différences de réalisation de la liaison apparaissent négligeables. Ils donnent comme exemple les contextes « + Nom pluriel » qui sont liés dans 80% des cas, indépendamment de leur longueur.

3.4.1.3 La longueur du groupe syntaxique auquel appartient M1

La prise en compte de la longueur du contexte syntagmatique auquel appartient M1 a elle aussi été signalée par Delattre (1955). Cependant, elle n'a pas été reprise par la suite. En affirmant que « *la longueur des éléments à lier peut avoir un effet notable. Plus le sujet est long, moins il unit avec le verbe qui le suit.* » (cf. 1955 : 47), il ne considère pas simplement la longueur du groupe syntaxique auquel appartient M2, mais également la longueur de celui auquel appartient M1.

La liaison se réalise d'autant plus que ce groupe est court. Aussi, dans l'exemple (26) où le groupe syntaxique auquel appartient M1 est plus long qu'en (25)⁵⁶, la probabilité de réalisation de la liaison serait plus faible :

(25) *Les enfants_attendent* : probabilité de réalisation de la liaison = **x**

(26) *Les plus petits des enfants_attendent* : probabilité de réalisation de la liaison < **x**

3.4.1.4 Bilan

Au vu des travaux exposés et des hypothèses avancées, on peut établir le bilan suivant :

- il paraît difficile de choisir entre les trois hypothèses : elles ont été inégalement testées à la fois quand on compare le nombre d'études et le nombre de contextes / mots envisagés ;

⁵⁶ Les exemples (25) et (26) sont repris de Delattre (1955 : 46).

- dans tous les cas, tester ces hypothèses demandent (i) de disposer de grands corpus (étant donné le problème de la variation, il faut que le corpus soit très important pour espérer disposer de différences significatives) ; (ii) de recourir à des outils sophistiqués permettant de classer les contextes (en faisant varier la cohésion entre les groupes, en faisant varier les termes lexicaux) ;

- à la différence de la syntaxe qui, certes ne détermine pas à elle seule la répartition de tous les contextes entre catégoriques et variables, mais qui en détermine au moins un (le contexte « Déterminant + Nom », cf. 3.2, pp. 57 et sq.), la longueur apparaît n'être qu'un facteur d'influence parmi d'autres ; elle ne détermine aucune répartition claire entre contextes catégoriques et contextes variables. Au mieux, elle permet d'établir une tendance parmi les contextes variables. Plus les expressions⁵⁷ en présence sont courtes, plus la liaison aura tendance à se réaliser. Dans tous les cas, il s'agit bien d'une tendance. Car on peut tomber sur des contre-exemples (mots « longs » liaisonnants), d'autant plus qu'il s'agit de contextes variables (justement soumis à différents facteurs d'influence).

Devant ces résultats somme toute peu discriminants, et à l'occasion de travaux cherchant à rendre compte des contextes variables, un certain nombre d'auteurs (cf. Ågren, 1973 ; De Jong, 1994, par exemple) se sont penchés sur la question de la fréquence lexicale des lexèmes. C'est ce que nous allons maintenant examiner.

3.4.2 La fréquence

L'hypothèse selon laquelle la fréquence lexicale des unités en contexte de liaison a un impact sur la production de la liaison, n'est pas nouvelle.

Ågren (1973) considère que ce sont les syntagmes les plus souvent employés qui réalisent le plus la liaison. Il observe que les liaisons sont plus souvent réalisées avec des M1 fréquents qu'avec des M1 rares. La liaison étant une « *survivance d'une prononciation ancienne* » (1973 : 14), certaines de ses formes ont été « *continuellement*

⁵⁷ Nous employons « expressions » à dessein, n'ayant pas tranché entre les trois propositions qui ont été examinées.

actualisées jusqu'à une quasi-mécanisation ». Il rappelle d'ailleurs que Langlard (1928) est un des premiers à faire état de la corrélation entre la fréquence d'emploi d'un terme et la fréquence de la liaison. Toutefois, il considère que tester cette hypothèse s'avère particulièrement compliqué et renonce au projet. Pour lui, il est impossible d'analyser la fréquence d'emploi du fait qu'elle interagit avec d'autres facteurs :

Il est cependant hors de ma portée de pousser plus loin l'examen de ce facteur et il faut constater que la *fréquence d'emploi* ne pourra pas être isolée des autres facteurs. Elle doit être en rapport étroit d'une part avec la *cohésion* (une forte cohésion suppose une *fréquence d'emploi* élevée), d'autre part avec la longueur des mots liants. (Ågren, 1973 : 28)

Dans cette section, nous présentons un certain nombre de travaux récents, psycholinguistiques et acquisitionnistes, qui ont relevé le défi de déterminer en quoi la fréquence lexicale pouvait être un paramètre déterminant pour la réalisation de la liaison.

Deux questionnements ont été essentiellement travaillés :

(i) La fréquence de la liaison est-elle corrélée à la fréquence lexicale de M1 et M2 ? Autrement dit, si la proposition « plus M1 et / ou M2 sont fréquents, plus la liaison a de chances d'être réalisée » est vraie ;

(ii) La fréquence est-elle corrélée à la fréquence de la co-occurrence M1 + M2 ? Autrement dit, si la proposition « plus la séquence M1 + M2 est fréquente, plus la liaison a de chances d'être réalisée » est vraie.

3.4.2.1 Fréquence de M1 et de M2

Récemment, une étude de Adda *et al.* (1999) menée sur un corpus de 155h a montré que les taux de réalisation de la liaison diffèrent selon que l'on considère ou non la fréquence lexicale des unités en jeu. Sur l'ensemble du corpus, le taux de réalisation de la liaison est de 55 %. En réduisant l'échantillon aux unités lexicales M1 et M2 les plus fréquentes, le taux de réalisation de la liaison s'élève à 64 %. Si l'on réduit encore l'échantillon aux M1 les plus fréquents, le taux de réalisation de la liaison culmine alors à 75 %. Ces résultats semblent indiquer que la fréquence de réalisation de la liaison est surtout dépendante de la fréquence lexicale de M1.

Ces résultats sont confirmés par d'autres analyses menées dans un cadre psycholinguistique. A l'aide d'un corpus lu et d'un corpus spontané, Fougeron *et al.* (2001a) mettent en avant l'existence d'une corrélation significative entre la fréquence de M1 et le taux de réalisation de la liaison. La liaison est plus souvent réalisée avec les M1 fréquents. Fougeron *et al.* ne sont toutefois pas aussi catégoriques que Adda *et al.* L'examen de la longueur de M1 et de la longueur du groupe syntaxique auquel appartient M1 révèle que, si en effet la fréquence lexicale de M1 est un facteur de réalisation, elle est également étroitement corrélée à la longueur de M1 (loi de Zipf, 1949) et au contexte syntaxique auquel il appartient (« contrainte de c-command »). Une autre étude (Fougeron *et al.*, 2001b) en arrive aux mêmes conclusions : les M1 les plus fréquents sont généralement les M1 les plus courts. Aussi, sachant que la liaison est plus souvent réalisée avec des lexèmes monosyllabiques (*cf.* point précédent), il s'avère difficile de déterminer qui de la longueur ou de la fréquence détermine le plus la liaison.

Par contre, les résultats de Fougeron *et al.* (2001a et b) indiquent que ni la fréquence, ni la longueur de M2 n'influencent la réalisation de la liaison : les M2 les moins fréquents et les plus longs ne sont pas moins liés que les M2 les plus fréquents et les M2 les plus courts (70 % de liaisons réalisées dans des contextes tel que « *gros actionnaire* », par exemple).

3.4.2.2 Fréquence de la co-occurrence

Dans une conception lexicaliste pure de la liaison, chaque occurrence est stockée avec son contexte de réalisation (*cf.* Bybee, 2001). La consonne de liaison est une consonne comme les autres, stockée dans le lexique avec des termes co-occurents. Ces unités lexicales sont stockées exhaustivement (par ex. *bon ami, bon enfant, des amis, mon ami, son petit ami, tous ses enfants, etc.*). La liaison n'est donc pas soumise aux contraintes morphosyntaxiques et prosodiques décrites par la phonologie abstraite, qui formalise spécifiquement le phénomène de la liaison comme segment flottant, épenthétique ou rattaché à M2 (*cf.* 4.5.1, 4.5.2 et 4.5.3 pour un bilan).

Cette dépendance de la liaison vis-à-vis de la fréquence lexicale de la séquence stockée, déjà notée par Long (1978) et Klausenburger (1984) notamment, a été étudiée par Fougeron *et al.* (2001b). Leurs résultats confirment cette hypothèse et montrent que

la fréquence de la co-occurrence M1 + M2 peut constituer un facteur statistiquement déterminant pour la réalisation de la liaison :

La fréquence de la séquence seule est donc importante, indépendamment de la fréquence du lieu. Les liaisons apparaissent donc plus fréquemment lorsque les lieux et liés sont des mots fréquents et quand ils forment une séquence de mots fréquents dans la langue. (Fougeron *et al.*, 2001b : 9)

3.5 Bilan sur les dimensions d'analyse

L'objectif de ce chapitre consistait à expliciter chacune des dimensions en jeu dans le phénomène de la liaison telles qu'elles ont été discutées dans la littérature et à mettre en avant son caractère variationnel.

Il est apparu que l'hypothèse la plus forte stipule que la liaison est avant tout conditionnée par la phonologie et la syntaxe qui permettent de spécifier et de classer les contextes dans lesquels elle peut apparaître. Toutefois, ces propositions n'ont aucune valeur explicative de la variation qu'elle présente et surtout, ne permettent pas d'étudier conjointement les autres dimensions en jeu dans la liaison.

Une analyse du phénomène à travers les dimensions sociolinguistiques et lexicales permet d'identifier des tendances :

- la liaison est un différenciateur social et un phénomène sociolinguistique inversé : (i) les locuteurs dont le niveau socioculturel est le plus élevé sont les plus à même de produire le plus de liaisons ; et (ii) plus le discours est soutenu, plus la liaison a tendance à être réalisée.

- la longueur et la fréquence des termes à lier semblent influencer la réalisation de la liaison : ce sont les mots les plus courts et les plus fréquents qui liaisonnent le plus.

Au bilan, il apparaît qu'aucune répartition claire entre les différents contextes de liaisons n'est déterminée et que l'articulation des différentes dimensions reste à faire.

Chapitre 4

Un processus phonologique

Dans ce chapitre, nous nous attachons à décrire les différentes propositions formelles qui ont été avancées par la phonologie, pour rendre compte du fonctionnement de la liaison. Ces travaux, qui trouvent tous leur origine dans la thèse de Schane (1965), ont été importants aussi bien sur le plan quantitatif que qualitatif, à tel point que pendant longtemps, on a pu « croire » que la liaison était un phénomène essentiellement phonologique. Aussi, parmi les travaux portant sur les différentes dimensions en jeu dans le phénomène, les propositions issues de la phonologie méritent qu'on leur accorde une attention particulière.

D'un point de vue phonologique, la liaison constitue, en français l'un des cas de réalisation de l'alternance consonne / zéro ($C \sim 0$). En français, on dénombre trois types caractéristiques d'alternance consonne / zéro : la flexion (verbale ou adjectivale), la dérivation (nominale ou adjectivale) et la liaison.

Le tableau suivant (Figure 4-1) donne un aperçu de ces trois cas d'alternance :

	Flexion	Dérivation	Liaison
t	<i>sort(t) vs sortent</i>	<i>peti(t) vs petite</i>	<i>peti(t) garçon vs petit_enfant</i>
z	<i>affreu(x) vs affreuse</i>	<i>dérider vs désobéir</i>	<i>de(s) galeux vs des_arnaques</i>

Figure 4-1 Alternances Consonne ~ Zéro

Dans le cas de la liaison, une même unité lexicale (présentant une consonne non fixe en finale) peut manifester une consonne alternante (*i.e.* réalisée ou non), le plus souvent selon l'initiale du mot qui lui succède : par exemple dans les séquences suivantes *ces*

enfants et *ces petits*, *ces* présente une alternance entre consonne et zéro ([sezãfã] vs [sepəti]).

Le traitement formel de l'alternance consonne / zéro est l'objet de nombreuses réflexions, lesquelles ont abouti à la mise en place de dispositifs formels que nous exposons dans les premières sections de ce chapitre. Ces dispositifs ne bénéficieront pas ici d'un traitement exhaustif. Nous nous contenterons de présenter ceux qui nous apparaissent comme étant (i) les plus représentatifs (approches générativistes classique (cf. 4.1) et multilinéaire (cf. 4.2)) et (ii) les plus actuels (Théorie de l'Optimalité (cf. 4.3) et Grammaire d'usages (cf. 4.4)). Nous nous appuyerons largement sur les différents comptes rendus récapitulatifs disponibles (Klausenburger, 1984 ; Encrevé ; 1988 et Tranel, 1995a et 1995b) et utiliserons en outre des travaux menés plus récemment (Tranel, 2000 ; Côté, 2003 ; Bybee, 2001 et 2003).

Chacune des propositions représente des formalismes sous formes de règles, de représentations ou de contraintes, et font émerger un problème central : le statut lexical de la consonne de liaison, qui fera l'objet d'une dernière section (cf. 4.5).

4.1 L'approche générative linéaire et les règles

L'approche générativiste classique se caractérise par la mise en place d'un formalisme à partir duquel sont construites de nombreuses règles.

4.1.1 La troncation

La première règle de liaison dans la théorie générative classique a été impulsée par Schane dans sa thèse de doctorat en 1965 (1965, 1967 et 1968). Il propose avec la *French Truncation Rule* (règle de troncation)⁵⁸, une tentative de formalisation de la liaison en soulignant le rapport existant entre la liaison et l'élision (disparition d'une

⁵⁸ La *French Truncation Rule* sera par la suite reformulée par différents auteurs tels que Milner (1967), Chomsky et Halle (1968) et Dell (1970).

voyelle finale devant voyelle), qu'il considère comme les deux aspects d'un même phénomène : la troncation. La règle est formalisée de la façon suivante⁵⁹ :

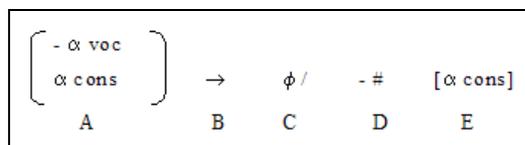


Figure 4-2 French Truncation Rule (Schane, 1968 : 10)

Schane considère qu'une même règle, généralisatrice, permet de traiter à la fois de la liaison et de l'élision, rapprochant ainsi deux phénomènes qui étaient jusqu'alors traités comme disjoints. La *French Truncation Rule* repose sur trois principes d'application que nous exposons ci-dessous.

1. Les deux processus (liaison et élision) sont corrélés et sont les deux faces d'un même processus

[...] a final vowel is deleted or truncated before another word beginning with a vowel, whereas a final consonant is deleted before another word beginning with a consonant. In order to be neutral between the terms *elision* and *absence of liaison*, we shall often refer to this one and the same process as *truncation*. (Schane, 1968 : 2)

L'élision est la chute d'une voyelle devant voyelle ; la liaison est la chute d'une consonne devant consonne. Autrement dit, la consonne de liaison est présente dans la forme phonologique sous-jacente. Pour Schane, ce qu'il convient d'expliquer, c'est la disparition de cette consonne et non son apparition. Le processus dans son entier concerne à la fois les consonnes et les voyelles, d'où le nom de « règle de troncation »⁶⁰.

2. Les segments phonologiques se répartissent en classes naturelles suivant leur comportement à l'égard de la troncation

⁵⁹ Cette règle se lit de la manière suivante : (A) une voyelle (- α) ou une consonne (+ α), (B) s'interprète, (C) comme l'élément nul (autrement dit, l'élément en question chute), (D) si elle est placée en finale de mot (devant une frontière #), (E) et si après cette frontière on trouve un élément de même nature (c'est-à-dire une voyelle (- α) ou une consonne (+ α)).

⁶⁰ C'est ainsi que le terme « *truncation* » a longtemps masqué le fait que la règle en question traitait en réalité de l'élision et de la non-liaison.

Ainsi :

Les voyelles (V)	- tombent devant voyelles (<i>l'ami</i>) et semi-voyelles (<i>l'oiseau</i>) - et se maintiennent devant consonnes (<i>le manteau</i>) et liquides [l] et [r] (<i>le lièvre</i>)
Les semi-voyelles (S)	- entraînent la chute de la voyelle précédente - et ne sont pas tronquées en fin de mot (<i>un pareil maladroit</i>)
Les liquides (L)	- entraînent la chute de la consonne précédente - et ne sont pas tronquées en fin de mot (<i>un pareil maladroit</i>)
Les consonnes (C)	- chutent devant consonnes (<i>un grand pouvoir</i>) - et se maintiennent devant voyelles (<i>un petit ennui</i>) - les occlusives finales sonores sont assourdis en contexte de liaison. Par exemple, le /d/ final de <i>grand</i> se réalise /t/ en liaison (devant voyelle). Les spirantes sourdes sont quant à elles voisées en contexte de liaison. Enfin, les liquides entraînent la chute de la consonne précédente et ne sont pas tronquées en fin de mot (par ex. <i>il pleure</i>)

Figure 4-3 Répartition des segments en classes naturelles

Ce principe peut lui-même se résumer en deux règles⁶¹ : la première porte sur les voyelles (Figure 4-4), qui chutent lorsqu'elles précèdent une frontière de mot suivie d'une voyelle (V) ou d'une semi-voyelle (S), la seconde porte sur les consonnes (Figure 4-5), qui peuvent elles aussi chuter si elles précèdent une frontière de mot suivie d'une consonne ou d'une liquide :



Figure 4-4 Chute de la voyelle

et :



Figure 4-5 Chute de la consonne

⁶¹ Conformément au système de notation adopté par SPE (1968), # désigne une limite d'unité, représentée par un blanc dans la topographie usuelle, et ## une pause forte entre deux unités consécutives.

Ces deux règles correspondent à la règle de troncation citée précédemment, à ceci près que les environnements conditionnant la chute de la voyelle ou de la consonne sont affinés, grâce à la désignation des classes naturelles (V, C, L, S) responsables du phénomène.

3. Les règles de troncation et de chute de consonnes finales sont cycliques

Pour une séquence de structures emboîtées les unes dans les autres et définies par l'analyse syntaxique, ces règles s'appliquent plusieurs fois, des unités les plus petites aux plus grandes. Pour une séquence de type [GN [ART *de* + S] [N *ami* + S] [ADJ *alemã* + S] GN], les deux règles s'appliqueront simultanément aux unités ART, ADJ et GN, puis à l'unité immédiatement supérieure, GN. Ce principe s'appuie sur une conception théorique générale. Dans le cadre choisi par Schane, celui de *The Sound Pattern of English*⁶², le rapport entre règles cycliques et règles non cycliques est déterminé de la façon suivante :

1. toutes les règles non cycliques s'appliquent à l'intérieur du mot phonologique ;
2. seules les règles cycliques peuvent s'appliquer au-delà du niveau de ce mot ;
3. aucune règle non cyclique ne peut suivre une règle cyclique.

La règle de troncation réussit à rendre compte de façon satisfaisante d'un grand nombre de phénomènes syntactiques du français et une même règle, permet de traiter à la fois des phénomènes de l'élision et de la liaison. Son importance englobe la théorie phonologique dans son ensemble puisqu'elle répartit l'ensemble des phonèmes du français en quatre classes naturelles (Voyelles, Consonnes, Semi-voyelles et Liquides).

Comme nous l'avons déjà signalé, le schéma que propose Schane traite essentiellement de l'absence de la liaison et de l'élision. Prenons un exemple de dérivation qui permet d'illustrer l'application de la *French Truncation Rule*.

Dans le tableau ci-dessous (Figure 4-6), l'exemple de *trop facile* montre l'effet de la *French Truncation Rule* et l'exemple de *trop aisé*, parce qu'il ne fait pas entrer en jeu la *French Truncation Rule*, fait apparaître un cas de liaison :

⁶² Chomsky N. et Halle M. (1968). *The Sound Pattern of English*. New York : Harper and Row ; trad. fr. Encrevé, P. *Principes de Phonologie générative*. Paris : Le Seuil, 1973.

Forme sous-jacente	/((tR op#)Adv (fasil#)sAdj)sAdj	/((tR op#)Adv (eze#)sAdj)sAdj
Effacement des parenthèses internes	(tR op#fasil#)sAdj	(tR op#eze#)sAdj
Règle de Troncation	(tR o#fasil#)sAdj	La règle ne s'applique pas
Effacement final des parenthèses	tR o#fasil#	tR op#eze#

Figure 4-6 Application de la French Truncation Rule

La *French Truncation Rule* ne traite donc pas de la liaison, mais bien de son absence :

Then elision for vowels and absence of liaison for consonants can be considered as the same process: deletion of a segment in word final position. (Schane, 1968 : 2)

Pour Encrevé (1988), la méthode utilisée par Schane se justifie par le fait que ce dernier rend compte de la grammaire d'une langue qui lui est seconde :

[...] la liaison n'est traitée par sa grammaire que dans les cas où elle *n'a pas lieu* [...] la règle qu'il pose consiste à *supprimer* ces consonnes dans les cas de non-liaison : ce qui revient à faire passer de l'écriture à la prononciation effective. [...] Schane suit d'autant plus facilement cette voie que c'est celle que lui proposait l'apprentissage de la forme phonique d'une langue étrangère à partir d'une graphie de cette langue [...]. (Encrevé, 1988 : 89)

Cette règle de troncation qui lie deux phénomènes (liaison et élision) oblige à mettre de côté les problèmes posés par la liaison facultative, et justifie le recours à une Règle d'effacement de la consonne finale (*Deletion Rule*). Cette dernière règle (Figure 4-7), qui permet de rendre compte de l'absence de la consonne, dérive de la forme des mots au niveau sous-jacent et s'applique comme suit (cf. Schane, 1968 : 13 ; Encrevé, 1988 : 92) :

Obligatoirement	- en position finale de syntagme - dans un nom singulier
Facultativement	- dans un nom pluriel

Figure 4-7 Règle d'effacement de la consonne finale

Cette nouvelle règle d'effacement permet à Schane de réintroduire le lien entre syntaxe et phonologie, tout en traitant de la variation qui ne trouve pourtant aucune place dans le cadre qu'il propose⁶³.

La règle de troncation proposée par Schane est soutenue par Milner (1967) qui propose une Règle de Réajustement des Frontières portant sur le nombre et la nature des frontières insérées dans la structure de surface par les composants syntaxiques. Ce principe de réajustement des frontières repose sur trois postulats (Encrevé, 1988 : 92) :

1. Les éléments de trait [-seg]⁶⁴ se rencontrent dans les suites phonologiques où ils sont insérés, en correspondance avec la parenthésisation des structures syntaxiques de surface ;
2. ces frontières sont un moyen de représenter l'information syntaxique par laquelle certains phénomènes phonologiques semblent déterminés ;
3. cette information doit être réajustée pour tenir compte de la différence entre les constituants proprement syntaxiques et ceux où s'exercent les règles phonologiques.

Cependant, la règle de troncation ne peut régler le problème que pose la troncation d'une consonne devant consonne et la troncation d'une consonne devant pause. Schane (1974) règle définitivement la question de la chute d'une consonne devant consonne et devant pause, en abandonnant la *French Truncation Rule* et la *Deletion Rule* proposées dans les ouvrages de 1965 et 1968. Il propose alors une Règle unique de suppression de la consonne, qui a la forme suivante :

$$C \rightarrow \emptyset / _ \text{---} ([\text{-seg}]) \{C, \#\}$$

Figure 4-8 Règle unique de suppression de la consonne

Schane aura donc traité de l'élision et des cas où il n'y a pas liaison, comme le constate Encrevé (1988 : 91), qui justifie la persistance de cette règle dans les écrits des générativistes influencés par la norme orthographique de la façon suivante :

⁶³ Voir Pagliano (2003 : 214).

⁶⁴ Autrement dit, les frontières.

C'est la force de cette tradition [orthographique] qui explique que Schane ait pu, malgré lui, faire prendre cette Règle de troncation pour une Règle de liaison, même par des phonologues francophones natifs pour qui l'expérience des mots n'était pourtant que seconde.

La règle de troncation pose différents problèmes qui sont largement décrits par Encrevé (1988). D'une part, traiter ensemble les deux phénomènes que sont l'élision et la liaison ne peut se faire « *qu'à condition de ne traiter que les cas où la non-liaison est obligatoire comme l'est l'élision* » (Encrevé, 1988 : 90). Traiter de la variation dans un modèle de ce type est donc impossible. D'autre part, le formalisme proposé illustre la « dépendance » de Schane à l'orthographe, qui ne lui permet pas de rendre compte de :

[...] ce qu'a de spécifique la consonne de liaison en français à savoir son statut très particulier, qui fait qu'elle peut pour le même mot, être présente ou absente sans que l'identité du mot soit affecté » (Encrevé, 1988 : 91)

4.1.2 La métathèse

Dell (1970) rejette l'unification de la liaison et de l'élision proposée par Schane (1965, 1968)⁶⁵. Il propose alors une Règle de Liaison, qu'il pose comme obligatoire. Cette règle est formalisée ainsi :

LIAIS (obl) :	[-syll]	#	[+syll]	=	2	1	3
	1	2	3		2	1	3

Figure 4-9 Règle de Liaison

La règle de liaison déplace obligatoirement un segment non vocalique à l'initiale du mot suivant si celui-ci commence par une voyelle (le trait [+syll] caractérisant les voyelles). Elle permet, par exemple, de rendre compte de l'homophonie de *petit homme* et *petit tome*. Comme l'indique Encrevé (1988), la Règle de Liaison de Dell constitue en

⁶⁵ Selkirk (1972) suit Dell quant au rejet de la règle de troncation de Schane. L'un comme l'autre, dans des perspectives toutefois opposées puisque Selkirk rejette également la règle de liaison de Dell, suivent Milner (1967) et sa proposition de réajustement des frontières morphologiques.

fait une règle de resyllabation et, par conséquent, traite plus de l'enchaînement que de la liaison proprement dite. Avec la Règle de Liaison, Dell adopte le format transformationnel proposé par *SPE* pour traiter les métathèses, où il y a échange de position entre deux éléments de même nature. Mais il confond métathèse et resyllabation, car il y a modification de l'appartenance d'un ou plusieurs segments à une structure syllabique, sans aucune métathèse⁶⁶.

Cette Règle de Liaison est suivie d'une Règle de Troncation qui efface toute séquence d'obstruante en fin de mot et donc, toute consonne qui ne se trouve pas en contexte de liaison (Encrevé, 1988 : 25) :

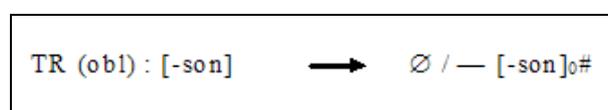


Figure 4-10 Règle de Troncation

Nous proposons ici un exemple d'application des deux règles précédemment énoncées. Cet exemple montre pourquoi il y a homophonie entre *petit homme* et *petit tome*.

	<i>petit homme</i>			<i>petit tome</i>		
Forme sous-jacente	pəti	#	ɔm	pəti	#	tɔm
	1	2	3	1	2	3
LIAIS (obl)	pəti#	t	ɔm	Ne s'applique pas		
	2	1	3			
TR (obl)	Ne s'applique pas			pəti	#	tɔm
	1	2	3	1	2	3
Surface	pəti	#	tɔm	pəti	#	tɔm
	2	1	3	1	2	3

Figure 4-11 Application des Règles de Liaison et de Troncation

⁶⁶ Encrevé (1988 : 99) indique que le terme de *métathèse* utilisé par Dell est inadéquat et devrait être remplacé par celui de *resyllabation*. La métathèse correspond à l'« échange de position entre deux éléments de même nature – ce qui n'est pas le cas pour les éléments en jeu dans sa règle : frontière de mot et CL » et la resyllabation à la « modification de l'appartenance syllabique d'un ou plusieurs segments sans aucune métathèse ».

Encrevé (1988 : 25-26) souligne que le problème posé par ces représentations est qu'en échappant à la Règle de Troncation, elles échappent également à LIAIS (cf. Figure 4-9). LIAIS ne constitue en fait qu'une règle d'enchaînement et non une règle de liaison. Dell traite le maintien de CL et sa resyllabation comme un seul et même phénomène. De plus, le système proposé par Dell entraîne l'impossibilité d'enchaînement pour une consonne précédant une CL lorsque la liaison n'est pas effectuée (par ex. *il part à Paris* [ilpar∩apari]).

Dans ses travaux ultérieurs, Dell (1973) abandonne sa Règle de Liaison et pose une Règle de Troncation équivalente à celle de Schane, que nous avons exposée précédemment (cf. Figure 4-2). En outre, Dell pose que la liaison est un phénomène résultant de deux conditions (phonologique et syntaxique) interdépendantes :

[...] il faut que deux conditions soient remplies simultanément. Il y a d'une part une condition d'ordre phonologique qui stipule que le mot à consonne latente doit être suivi d'un mot commençant par une voyelle, et d'autre part une condition syntaxique qui stipule que le rapport syntaxique entre les deux mots doit être suffisamment étroit. (Dell, 1973 : 42)

La première condition, d'ordre phonologique, stipule que le mot à consonne latente doit être suivi d'un mot commençant par une voyelle. Par exemple, dans *vous écoutez* [vuzekute], le [z] final de *vous* est une consonne latente. La liaison peut donc se réaliser puisque cette consonne latente est placée devant une voyelle. En revanche, on ne prononce jamais cette consonne lorsque le mot auquel elle appartient se trouve en fin de phrase (par ex. *levez-vous* [løvevu]) ou précède un mot commençant par une consonne (par ex. *vous regardez* [vurəgarde]) (cf. 3.1.1).

Lorsque cette condition phonologique est remplie, intervient la seconde condition, laquelle indique que le rapport syntaxique entre les deux mots doit être suffisamment étroit (cf. 3.2.2).

4.1.3 L'épenthèse

La solution de l'épenthèse est supportée par les phonologues voulant traiter les formes phoniques indépendamment de l'orthographe (Klausenburger 1974, 1976, 1977, 1978a et 1978b ; Tranel 1976, 1978 et 1981 ; Kaye et Morin, 1976).

The reaction to the transformational-generative deletion analysis of Schane, Dell, [...] has been led, during the 70's, by Klausenburger and Tranel, who, working independently, arrived at similar conclusions concerning the flaws of the deletion account. (Klausenburger, 1984 : 18)

Le mouvement est initié par Klausenburger (1974) qui ne considère désormais plus la liaison comme relevant de la présence d'un segment sous-jacent. Il propose alors une nouvelle hypothèse : celle de l'épenthèse⁶⁷. Cette règle d'épenthèse que Klausenburger appelle *Liaison*, est de la forme suivante :

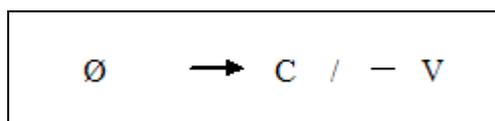


Figure 4-12 Règle d'épenthèse

La règle d'épenthèse indique que la consonne de liaison ne peut apparaître que grâce à la présence d'une voyelle. Selon Encrevé (1988), l'inconvénient d'une telle solution est de ne fournir aucune indication quant au statut de la CL. La consonne de liaison n'est effectivement pas représentée dans le lexique :

[...] le *t* de *petit ami* n'est inscrit en rien dans la représentation lexicale du mot *petit* (/pəti/), c'est-à-dire qu'il a le même statut que n'importe quelle autre consonne n'appartenant pas à ce mot, et que n'importe quel son n'appartenant pas à la langue française [...]. (Encrevé, 1988 : 105)

Parmi les autres phonologues qui adoptent la solution de l'épenthèse, Tranel (1974) fait figure de référence dans le sens où il fournit la formalisation la plus complète du

⁶⁷ Dans le premier chapitre de sa thèse, Pagliano (2003) définit l'épenthèse et ses caractéristiques. Elle précise que l'épenthèse est un « élément dont on ne trouve pas trace en diachronie » (2003 : 8), à caractère non étymologique, non sous-jacent et non prédictible par le contexte (2003 : 12).

processus. Contrairement à Klausenburger (1974), Tranel tente de préciser les mécanismes formels qu'implique une telle règle. Il formalise tout d'abord une règle d'insertion de *n* pour les liaisons des voyelles nasales ne présentant pas d'alternance de type \tilde{v} / VN (1981 : 240) :

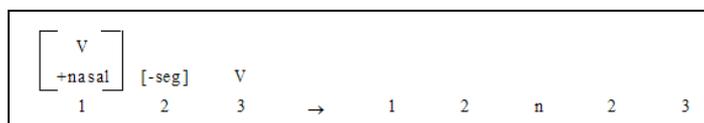


Figure 4-13 Règle d'insertion de *n*

Tranel (1976) ajuste la règle d'insertion de *n*, en revenant à une représentation sous-jacente des adjectifs (par ex. *bon* dont la représentation est désormais du type /b \tilde{o} /). Il accompagne cette règle d'une règle phonologique de dénasalisation de la voyelle formalisée ainsi :

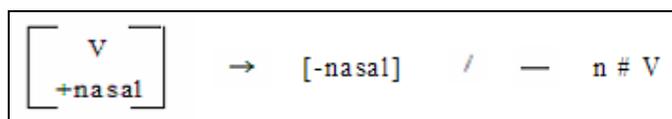


Figure 4-14 Règle de dénasalisation de la voyelle

Suite à ce travail sur les nasales, Tranel (1976) propose le même type de règle mais cette fois-ci relative au *z* du pluriel⁶⁸, contenant des informations syntaxiques tout en conservant l'emploi des frontières morphologiques :

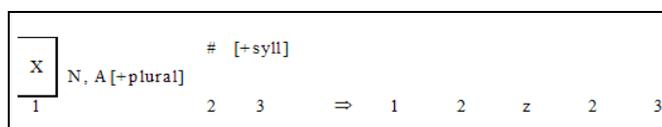


Figure 4-15 Règle d'insertion de *z*

En 1981, dans un ouvrage intitulé *Concretness in Generative Phonology, Evidence from French*, Tranel (1981a) présente l'ensemble de ses propositions (règles d'insertion

des consonnes de liaison *z*, *t*, *p* et *n* de mots invariables ; règles d'insertion du *z* du pluriel mais aussi règles d'insertion des consonnes de liaisons verbales *z*, *t* et *r*). L'idée défendue par Tranel est que les processus phonologiques sont liés à la cohésion syntaxique.

Le principal avantage de la représentation de ce type d'informations est que cela permet d'éviter le recours aux règles de réajustement des frontières (*cf.* Encrevé, 1988 : 120). Dans son analyse, Tranel rend compte des alternances consonne / zéro de la façon suivante (cité par Klausenburger, 1984 : 20) :

1. *Liaison with verbs*
 - a. *optional liaison with person endings*
 - b. *obligatory liaison with person endings*
 - c. *infinitive liaison (optional)*
2. *Liaison with masculine singular adjectives (obligatory)*
3. *Liaison with invariable words (optional)*

Mais, comme le fait remarquer Encrevé (1988 : 119), ce travail « *vient malheureusement bien trop tard* », les propositions faites par Tranel ne reconnaissant aucun rôle aux structures non segmentales. D'autre part, une solution par épenthèse ne permet en aucun cas de prédire la consonne de liaison et ne constitue en fait qu'un modèle non explicatif. Pagliano (2003) considère d'ailleurs que l'épenthèse est la solution la *moins satisfaisante* pour expliquer les alternances consonne / zéro de la liaison :

[...] la nature de la consonne sous intérêt est à spécifier à chaque fois, puisque dans la plupart des cas de liaison elle n'est pas prédictible par le contexte, ni phonologique, ni morphologique, ni syntaxique. De plus, l'appel à une règle d'insertion n'est pas explicatif mais uniquement représentationnel. Dans les cas où l'apparition de la consonne est liée à une contrainte syntaxique, *i.e.* quand elle apparaît à la liaison dans un contexte pluriel, l'épenthèse pourrait permettre de rendre compte de l'invariabilité de la nature de la consonne, mais rien n'autorise à privilégier cette solution à celle qui consiste à voir en la consonne intrusive une marque morphologique. (Pagliano, 2003: 245)

⁶⁸ Kaye et Morin (1976) considèrent également que la liaison est épenthétique. La seule exception concerne le *z* du pluriel qu'ils considèrent être un préfixe, un préfixe du pluriel pour les mots à initiale vocalique. (*cf.* Encrevé 1988 : 109).

Wauquier est encore plus sévère dans sa critique de l'épenthèse, arguant que :

[...] les motivations de la règle d'épenthèse ne paraissent pas moins opaques que celles de la règle de délétion et apparaissent même encore plus "anti-naturelles" pour rendre compte des données de compréhension. (Wauquier, 2005 : 39)

4.1.4 Bilan d'une approche par règles

Le traitement de la liaison par règles proposé par les phonologues générativistes classiques repose, pour reprendre les termes de Tranel (2000 : 44), sur « *une même constante : l'existence présumée de processus phonologiques d'effacement ou d'insertion opérant dans certains contextes* ». Le travail descriptif fourni est phénoménal. Cependant, une approche par règles, telles que celles qui viennent d'être décrites dans cette section, comporte de nombreux inconvénients. Par exemple, la perspective de Schane (1967), laquelle considère le phénomène de la liaison comme *unique et unitaire* et essentiellement phonologique⁶⁹, montre son incapacité à se dégager de l'orthographe et ne peut expliquer la tripartition des liaisons.

Tranel (2000), dans son article consacré à l'implémentation du schwa et de la liaison dans le cadre de la Théorie de l'Optimalité⁷⁰, justifie l'abandon des analyses par règles par le fait que celles-ci sont purement descriptives et sans grande valeur explicative des phénomènes phonologiques : « *les règles nous disent ce qui se passe, mais elles n'expliquent pas pourquoi les choses sont ainsi* » (Tranel, 2000 : 44). Le second argument présenté par Tranel est celui de l'insuffisance de la portée des analyses de ce type : elles ne peuvent « *couvrir les données de façon exhaustive* » (Tranel, 2000 : 46). Les règles sont incomplètes et nécessitent, pour la plupart, la mise en place de nouvelles règles, cette fois-ci de réajustement. Enfin, toujours selon Tranel (2000 : 47), la grammaire serait plus « *simple* » sans aucune règle ne modifiant les représentations sous-jacentes.

⁶⁹ A ce propos, voir Laks (2005a) et Turcsan (2005).

⁷⁰ Cf. 1.2.3.

4.2 L'approche multilinéaire et les représentations

La fin des années 70 est décisive pour la phonologie qui connaît un renouvellement théorique important. En opposition à un traitement de la liaison par règles et surtout en rupture avec SPE, un nouveau dispositif générativiste se développe à partir de 1976, avec la thèse de Goldsmith : la phonologie multilinéaire.

On sait que la théorie de *SPE* est considérée comme périmée depuis quelques années. Mais cette constatation s'est imposée en grande part à la faveur de l'élargissement de la base empirique de la phonologie générative, l'étude des langues tonales africaines et l'étude de la morphologie des langues sémitiques. [...] L'unilinéarité de *SPE* a été trouvée manifestement insuffisante pour rendre compte des faits phonologiques en cause. (Encrevé, 1988 : 159)

Dans ce nouveau cadre, on propose alors de remplacer le trait [vocalique], une qualité phonémique, par le trait [syllabique] qui désigne une position dans une architecture susceptible de constituer un pic de syllabe et ce, sans introduire aucune vision hiérarchique de la phonologie. L'objectif est de dissocier, entre autres, la mélodie des constituants squelettal et syllabique. Ce sera Kahn⁷¹ (1976) qui réintroduira cette notion de syllabe en phonologie générative. Conjointement, Goldsmith (1976), confronté aux problèmes des tons dans les langues africaines, introduit une nouvelle notation, dite *autosegmentale*, qui permet d'éviter les longs cycles de dérivations et d'assignation des accents propres à SPE (1968). Cette notation, la réintroduction de la syllabe, et la vision configurationnelle de la phonologie scelleront la naissance de la phonologie multilinéaire et ébranleront l'unilinéarisme transformationnel de SPE.

Cette nouvelle phonologie abandonne ainsi les matrices de traits utilisées par SPE (1968) : chaque trait distinctif est disposé sur un palier autosegmental. Certains principes, les conditions de bonne formation (WFC : Well Formedness Conditions), permettent d'obtenir des structures optimales, bien formées. Les deux principales WFC sont :

1. la LCC : *Line Crossing Constraint* qui interdit le croisement des lignes reliant les paliers autosegmentaux ;

⁷¹ Kahn (1976) dans sa thèse *Syllable Based Generalizations in English Phonology*.

2. l'OCP : *Obligatory Contour Principle* qui interdit l'adjacence d'éléments de même nature sur un même palier.

Encrevé (1988), dans son ouvrage, retrace l'histoire du traitement de la liaison dans les différents courants phonologiques antérieurs à la phonologie tridimensionnelle. D'autres linguistes, tel que Klausenburger (1984), ont déjà établi cette évolution, présentant les travaux les plus importants ayant permis l'évolution du traitement de cet objet linguistique qu'est la liaison.

Dans son ouvrage, Klausenburger (1984) opte pour « un état des lieux » définissant et résumant les différents travaux sur la question. Encrevé (1988), avec son ouvrage *La Liaison avec et sans enchaînement*, établit quant à lui, non seulement un bilan épistémologique en critiquant les points faibles des différentes hypothèses formulées, mais intègre également un nouveau schéma explicatif de la liaison en le testant à un corpus. En outre, il associe une analyse phonologique et une approche sociolinguistique. Il traite de la liaison en appliquant les concepts de la phonologie tridimensionnelle.

Encrevé remarque que les thèmes travaillés en phonologie autosegmentale n'arrivent pas encore à proposer de traitement unifié et se propose d'explicitier sa propre théorisation en se servant des propositions lui semblant les plus pertinentes. Son ouvrage apparaît alors comme le seul travail cumulatif sur la question. Son approche reste parallèle à celle de Delattre : tout en présentant le champ théorique qui va lui permettre d'analyser ses résultats, il fait l'inventaire de toutes les formes stables de la liaison, formes à la fois syntaxiques et phonologiques. Il tente une approche sociolinguistique, reconnaissant qu'une des particularités de la liaison est d'être avant tout un objet de variation. Pour expliciter cette variation et ses différents aspects, il entreprend une approche labovienne du problème : la formalisation ne peut se départir de la variation.

Dans une approche multilinéaire, les éléments phonologiques (mélodie, constituants squelettal et syllabique) sont ainsi :

- (i) dissociés et représentés sur des niveaux différents ;
- (ii) mis en relation par la ligne du squelette ;
- (iii) et contraints par des principes universels :

Une phonologie multilinéaire n'est pas *ipso facto* une garantie absolue contre l'approche « excessivement formelle » de *SPE* et notamment contre l'absence de contraintes sur les règles. Bien des phonologues continuent imperturbablement, dans le cadre considérablement enrichi de la phonologie multilinéaire, à se servir des possibilités d'écriture formelle de *SPE*, qu'ils cumulent avec les possibilités des configurations autosegmentales pour écrire encore et toujours des règles ordonnées, et cycliques si cela facilite les choses. [...] Plus que jamais, l'avertissement du chapitre 9 de *SPE* vaut : il ne suffit pas d'aligner les formules ; encore faudrait-il qu'elles aient quelques liens avec « certains principes qui régissent les langues humaines ». (Encrevé, 1988 : 159)

L'application du principe de multilinéarité (dissociation autonome de chacune des dimensions) au phénomène de la liaison favorise le débat autour du statut de la consonne de liaison, qui est désormais différent des autres consonnes ; elle est considérée comme flottante, non rattachée au niveau syllabique :

Dans l'approche multilinéaire, la CL est considérée comme un segment flottant, c'est-à-dire à la fois une coda flottante, sans association avec le plan autosegmental syllabique, et une consonne flottante, sans association avec le squelette sur la ligne autosegmentale des segments. Ces caractéristiques de la CL font d'elle une consonne qui appartient lexicalement au mot (Mot1) mais qui possède une représentation distincte des consonnes toujours réalisées. (Dugua, 2006 : 23)

C'est en 1981 que la consonne de liaison est pour la première fois représentée comme un segment flottant associé à une position consonantique mais pas à un nœud syllabique (cf. Clements et Keyser, 1981).

Pour que la liaison soit réalisée, il est nécessaire de passer par une règle de liaison qui associe toute position consonantique non rattachée à une syllabe à la syllabe suivante. La consonne, non rattachée à la position squelettale, n'est donc pas un constituant syllabique.

Deux propositions de traitement émergent de cette notion de « flottance » des consonnes de liaison :

1. soit les consonnes de liaison, qui « *ont une existence dans la représentation lexicale, sont rattachées à une position segmentale mais non associées, dans la tire syllabique, à une structure σ* » (Wauquier-Gravelines, 2005 : 40)⁷²;

⁷² Cf. Clements et Keyser (1981) et (1983).

2. soit elles ne sont attachées à rien, elles sont « *purement flottantes sans aucune association* » (Wauquier-Gravelines, 2005 : 41)⁷³.

La conception du flottement la plus communément admise est celle que propose Encrevé (1988) dont nous avons déjà donné une description en 3.1.3.1 et 3.1.3.2. La consonne de liaison appartient à l'unité lexicale mais n'est associée à aucune position, ni squelettale, ni syllabique. Son ancrage n'est possible que si une position squelettale est disponible (en attaque ou en coda).

C'est à partir de ce modèle théorique qu'Encrevé trouve des réponses aux problèmes posés par l'ensemble des faits phonologiques qui sont en cause dans la liaison en français, dans l'enchaînement et dans le non-enchaînement. Cet ensemble de faits permet en effet de mettre en valeur les propositions de ce cadre et d'instaurer certaines conventions universelles et certains paramètres du français. Son analyse présente en outre l'avantage de résoudre certains problèmes posés par la structure syllabique.

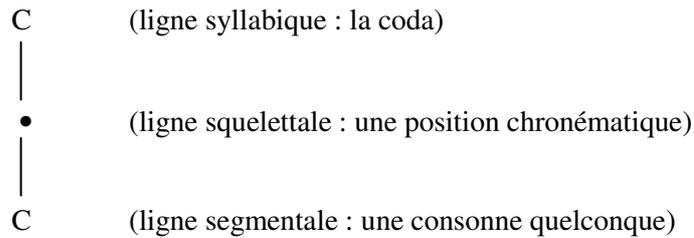
Encrevé présente la consonne de liaison en français comme une consonne qui appartient lexicalement au mot, tantôt réalisée tantôt non réalisée dans la prononciation.

Dans la représentation lexicale, il postule trois principes :

1. sur la ligne autosegmentale des segments, la CL est un segment flottant et donc sans association avec le squelette ;
2. dans le plan autosegmental syllabique, sur la ligne des constituants de la rime, une coda flottante, et donc sans association non plus avec le squelette, correspond avec ce segment flottant ;
3. sur la ligne du squelette, une position, sans aucune association, et disponible pour l'ancrage à la fois de la CL et d'un constituant syllabique.

La structure de la CL en français est donc représentée comme suit :

⁷³ Cf. Encrevé (1988).



Dans une structure de ce type, la CL appartient lexicalement au mot et a une représentation distincte des consonnes toujours réalisées dans la prononciation. Cette dernière dépend de cette CL, de l'opération d'ancrage ou non dans le squelette, ainsi que de l'interprétation syllabique de la position dans laquelle elle s'ancr.

Cependant, une coda flottante ne peut pas s'associer à n'importe quel constituant de la ligne A (attaque), R (rime). Pour y remédier, Encrevé postule une nouvelle *Convention universelle d'association des codas flottantes* :

[...] une coda flottante ne peut s'associer sur la ligne des constituants syllabiques A, R qu'avec l'unité R dominant le noyau qui la précède sur la ligne N, C. (Encrevé, 1988 : 171)

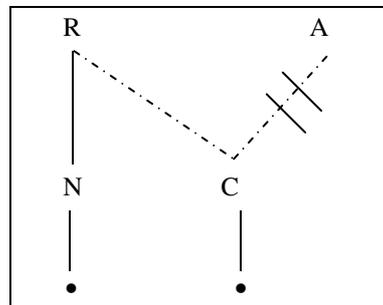


Figure 4-16 Convention universelle d'association des codas flottantes

Cette convention permet d'éviter une configuration mal formée (représentée par la ligne en pointillé barrée où un nœud A domine un nœud C).

La consonne finale flottante (CL) se distingue donc de la consonne finale fixe qui est, elle, rattachée à un point squelettal :

[...] les consonnes toujours prononcées que nous appelons consonnes fixes, CF, sont ancrées dans le squelette dès la représentation lexicale, qu'elles soient en position initiale, interne ou finale dans le mot. (Encrevé, 1988 : 173)

Cette analyse spécifique de la consonne de liaison et de la consonne finale fixe fait apparaître une structure du français particulière : à la finale de mot un constituant syllabique coda est toujours flottant (que ce soit une CL ou une CF). Il élargit même cette hypothèse, allant jusqu'à dire qu'en français, tout autosegment syllabique coda est flottant. Pour lui, en français, seules les consonnes situées à l'attaque d'une syllabe seraient, dans tous les cas, interprétées syllabiquement au niveau sous-jacent :

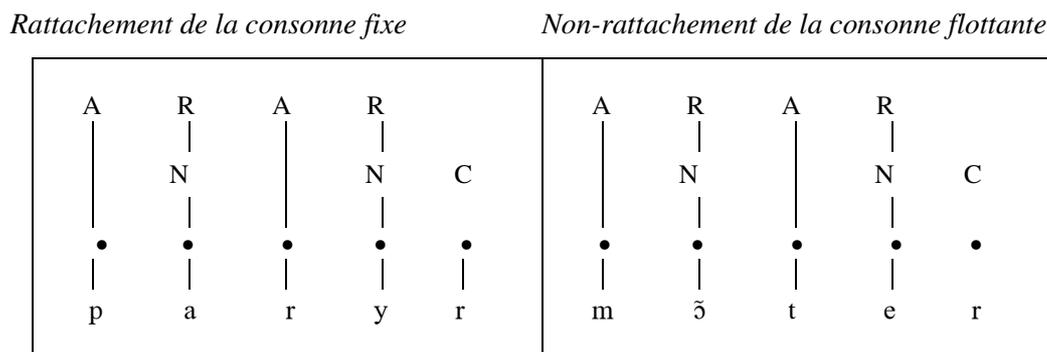


Figure 4-17 Représentation du (non)-rattachement des consonnes finales

Traiter le phénomène de la liaison par le biais d'une approche multilinéaire telle que celle qui a été proposée par Encrevé (1988) consiste à adopter une conception de la CL comme segment flottant et à considérer que la liaison n'est pas le résultat d'une série de règles mais celui d'une succession de conventions de bonne formation propre au français.

La principale critique faite à l'encontre d'une telle représentation est avancée par Pagliano (2003). L'association entre la consonne flottante et la position squelettale n'est pas automatique dans la représentation (*cf.* De Jong, 1988 ; Paradis et El Fenne, 1995) ; or, dans un cadre autosegmental, toute association d'un constituant mélodique dans une position squelettale l'est (*cf.* Vergnaud, 1982).

Par ailleurs, Wauquier-Gravelines (2005) rend compte du fait que dans la proposition d'Encrevé, la consonne de liaison est « doublement » flottante puisqu'elle est syllabiquement et segmentalement flottante. Elle n'a par conséquent aucune association :

Si l'on retient ce formalisme, la CL appartient bien à l'unité lexicale (ce qui explique ses caractéristiques segmentales) mais hors de toute ligne, dans une espèce de non linéarité ou de linéarité suspensive [...]. Elle est

donc « suspendue » en attente d'une position et non soumise à la segmentation qui s'applique aux lignes syllabiques et segmentales et qui permettent de dégager respectivement les consonnes « ordinaires » et les syllabes bien formées. Ce faisant, elle échappe d'une certaine manière à la séquentialité temporelle exprimée dans le signal qui lui imposerait une position *a priori*. Elle n'est donc pas traitée en temps réel de la même manière qu'une consonne ordinaire même si elle en a quasiment les caractéristiques acoustiques. (Wauquier-Gravelines, 2005 : 41)

Laks (2005a) considère quant à lui qu'il est nécessaire que les modèles phonologiques rendent compte du fait que les représentations lexicales sont influencées par l'orthographe. C'est l'effet Buben (*cf.* Chevrot et Malderez, 1999 ; Buben, 1935 ; Martinet, 1945) ; il fait consensus :

[...] nous proposons d'inscrire *l'identité visuelle du mot* dans la représentation cognitive en posant comme entrées lexicales des constructions autosegmentalisées comprenant, outre des lignes phonologiques, sémantiques et syntaxiques, une ligne autosegmentale de représentation graphique comprenant au moins la consonne finale et la limite de mot graphique car, pour maîtriser les liaisons, spécialement celles postposées à une catégorie principale qui sont facultatives et peu fréquentes, « il faut se représenter le mot écrit. » (Thimonnier, 1970 : 102). (Laks, 2005a : 115)

Laks (2005a : 119-120) rappelle qu'un certain nombre de consonnes aujourd'hui fixes ne l'ont pas toujours été (par ex. *lis, mais, jadis*). Ce qu'il appelle « *l'effet de fixation* » pointe l'influence de la graphie : « *c'est l'identité visuelle des mots et la lecture qui rétablit et fixe progressivement les finales* ». La position de Laks n'est donc pas de rejeter les représentations multilinéaires mais d'y encoder une nouvelle ligne autosegmentale, contenant l'information graphique (GRAPH). Cette ligne segmentale supplémentaire intégrerait, minimalement, les frontières de mots et la consonne finale.

4.3 La Théorie de l'Optimalité et la notion de contrainte

La Théorie de l'Optimalité⁷⁴ (Prince et Smolensky, 1993) s'inscrit dans un programme de recherche générativiste proche du générativisme linéaire, qui repose sur un certain nombre de principes résumés ici : les langues doivent satisfaire un certain nombre de contraintes universelles de bonne formation ; ces contraintes peuvent entrer en conflit à l'intérieur d'une langue particulière ; certaines contraintes sont violables ; et les contraintes sont ordonnables⁷⁵ :

D'un point de vue général, ce cadre se distingue par un rejet caractérisé du binarisme⁷⁶. [...]. Dans ce cadre, un « paramètre » n'est pas actif ou inactif : son rôle dans la grammaire sera d'autant plus important qu'il sera haut placé dans la hiérarchie. (Eychenne, 2006 : 52)

Le principe d'OT est de générer des contraintes universelles, applicables à toutes les langues, qu'il faut ensuite hiérarchiser pour chacune d'entre elles et pour chacun des phénomènes étudiés.

Dans la grammaire d'OT, les langues sont des systèmes intrinsèquement conflictuels, dans lesquelles s'opposent deux forces :

- les contraintes de marque, qui ne considèrent que l'output et favorisent les structures non marquées ;
- et les contraintes de fidélité, qui incitent au contraire les langues à conserver les oppositions lexicales.

Au lieu de concevoir les règles comme des montages s'appliquant à un input pour former un output par le biais de réajustements, OT génère (fonction « GEN ») un ensemble de candidats qu'elle évalue (fonction « EVAL ») en fonction des contraintes préalablement établies.

⁷⁴ *Optimality Theory*, désormais notée OT.

⁷⁵ Pour une description complète d'OT, cf. McCarthy et Prince, 1993a, 1993b, 1995 ; Kager, 1999 ; McCarthy, 2002. Voir aussi Boltanski (1999).

⁷⁶ Eychenne (2006) délaisse le binarisme d'une approche Principes et Paramètres, au profit de « l'approche gradiente de la grammaticalité que propose la théorie de l'optimalité » (2006 : 52). Il rejette également le binarisme de SPE intégré « tacitement » par OT, lui préférant l'approche « privative héritée

Concernant la liaison, Tranel (2000) est, à notre connaissance, le seul auteur à avoir proposé un traitement unifié de la liaison (et de l'élisio) dans ce cadre. C'est pourquoi cette section s'attachera à décrire ses propositions.

Tranel, sans prendre position par rapport à la nature lexicale de la consonne flottante, développe l'idée que la CL est différemment représentée dans la base par rapport aux consonnes dites « intrinsèques » :

[...] je ne prendrai pas position en ce qui concerne la nature de la source lexicale des consonnes de liaison. [Elle] n'appartient pas en propre au morphème [...] mais il s'agit, selon les analyses, ou bien d'une consonne finale flottante, ou bien d'une consonne empruntée au féminin, ou bien encore d'une consonne de déclinaison, en tout cas d'une consonne lexicalement disponible, mais non intrinsèque. (Tranel, 2000 : 51)

La consonne de liaison constitue un segment épenthétique lexicalement disponible, inséré lorsque les conditions syllabiques sont favorables.

La contrainte la plus forte proposée par Tranel est une contrainte anti-hiatus (*VV). Cette contrainte est en concurrence avec d'autres types de contraintes, telles que les contraintes DEP qui interdisent l'épenthèse (DEP (C) pour les consonnes et DEP (L) pour les voyelles). La CL ne se réalise généralement pas sous l'effet de la contraintes DEP (L), mais peut tout de même être insérée dès lors qu'il s'agit d'éviter un hiatus (*VV), la grammaire d'OT favorisant la domination de *VV sur DEP (L), soit : *VV >> DEP (L).

Une telle analyse présente plusieurs inconvénients.

Premièrement, pour Pagliano (2003 : 225-227), la syntaxe et la phonologie se situent à des niveaux différents et mêler leurs contraintes respectives n'est pas envisageable. L'analyse des cas de liaisons supplétives pose un problème majeur : pour deux allomorphes, l'un à finale vocalique et l'autre à finale consonantique, la hiérarchie va préférer le second, violant par là-même la contrainte d'accord en genre entre un déterminant et un nom (GENRE[SYN]), au profit d'une contrainte de type phonologique (*VV). Elle rappelle d'ailleurs que Tranel lui-même envisage des « solutions de

des phonologies particulières, intégrée dans une composante géométrique richement structurée » (2006 : 66-82).

résolution » : soit une sous-spécification du genre masculin, soit l'intervention de la contrainte –COD, soit enfin considérer que la contrainte portant sur le genre ne relève pas de la syntaxe mais de la morphologie.

Eychenne (2006) fait quant à lui remarquer que le manque de positionnement théorique de Tranel quant au statut lexical de la CL soulève un problème théorique majeur. Si la consonne doit être présente dans l'input et ce, quel que soit sa nature, il devient alors :

[...] difficile de concevoir comment l'apparition de la consonne de liaison en output peut violer une contrainte de la famille DEP qui demande à ce que tout élément de l'output ait un correspondant en input : si la consonne de liaison est présente en input elle ne peut pas causer la violation de DEP, au sens où cette contrainte est normalement comprise dans la Théorie de la correspondance. (Eychenne, 2006 : 353)

Pour Eychenne, soit le segment est présent dans l'input et son apparition en output ne viole pas DEP, soit le segment est absent en input et son apparition viole DEP. Par le biais d'une analyse de la liaison du pluriel, il démontre l'incapacité de l'analyse de Tranel à générer le bon candidat.

Pour conclure sur OT mais également pour introduire le dernier dispositif formel que nous exposons dans cette section, on note que pour Bybee (2001), le défaut d'OT est de séparer le lexique et la grammaire.

4.4 Théories basées sur l'usage

Dans les approches phonologiques par règles, par représentations ou par contraintes, les représentations sont décrites comme statiques et ne permettent pas la modification par l'usage ni des règles, ni des formes lexicales d'une langue⁷⁷ :

⁷⁷ Cf. Bybee (2001 : 14) : « [...] generative theories assume that the lexicon is a static list, and neither the rules nor the lexicon forms of a language are changed at all by instances of use. Similarly, as Pierrehumbert (1999) points out, all versions of Optimality Theory (Hayes, 1999 ; Prince and

[...] c'est la structure du lexique, la nature, la forme et le statut des représentations lexicales postulées qui gagent l'ensemble de l'analyse. [...] Même dans les approches représentationnelles, la représentation ne vaut que par les processus qu'elle motive. (Laks, 2005a : 103)

A *contrario*, les modèles exemplaristes (cf. Bybee, 2001 ; Pierrehumbert, 2002) abandonnent les systèmes abstraits et supposent que les représentations lexicales sont dépendantes de l'usage et que des schémas (*généralisations*) émergent de celles-ci :

Generalizations over forms are not separate from the stored representation of forms but emerge directly from them. In Langacker's terms, there is « *no rule/list separation* ». (Bybee, 2001 : 7)

La phonologie exemplariste s'inscrit dans le champ de la linguistique cognitive et se fixe comme objectif d'identifier et de déterminer le rôle du lexique, sa nature formelle et cognitive. Comme le souligne Bybee (2001), les modèles exemplaristes ont « *une conception procédurale de la phonologie* » (cf. également Wauquier-Gravelines, 2005 : 129) puisque :

[...] the lexical meaning may be propositional [...], but [...] the rest of language, surely all of the grammar and phonology is procedural knowledge. (cf. Bybee, 2001a : 40)

L'idée défendue par les modèles exemplaristes est que les expériences productive et perceptive, notamment « *the creative role of repetition* »⁷⁸, exercent une influence sur les représentations.

Le caractère plus ou moins répétitif et donc plus ou moins cumulatif de ces expériences suppose un stockage des représentations phonologiques (*types*) fondé sur la fréquence en contexte des réalisations de surface (*tokens*) :

In an exemplar model, each category is represented in memory by a large cloud of remembered tokens of that category. These memories are organized in a cognitive map, so that memories of highly similar instances are close to each other and memories of dissimilar instances are far apart. The remembered tokens display the range of variation that is exhibited in the physical manifestations of the category. (Pierrehumbert, 2000 : 3)

Smolensky, 1993, 1997) posit a strict separation of lexicon and grammar that makes it impossible to describe any of the interactions of phonology with the lexicon that are attested in the literature [...]. »

⁷⁸ Pour une description, voir Bybee (2001 : 8-10).

Cependant, les unités de base ne sont pas simplement des mots mais peuvent également être des suites de mots employées fréquemment et des formes figées (cf. Bybee, 2005b). Ce qui est le cas, par exemple, des suites d'unités impliquant une liaison. Ces différentes unités, ou constructions, constituent alors l'association stabilisée d'une forme et d'un sens, et l'on fait l'hypothèse qu'elles sont stockées en mémoire. Elles sont vues comme les unités de base de la grammaire et trouvent, dans les Grammaires de Construction (cf. Langacker, 1991 ; Goldberg, 1995 ; Croft, 2001), un nouveau cadre formel d'expression.

Dans les Grammaires de Construction, il est postulé que les niveaux sémantiques, syntaxiques et morphologiques interagissent :

In Construction Grammar, no strict division is assumed between the lexicon and syntax. Lexical constructions and syntactic constructions differ in internal complexity, and also in the extent to which phonological form is specified, but both lexical and syntactic constructions are essentially the same type of declaratively represented data structure : both pair form with meaning.» (Goldberg, 1995 : 7)

Le modèle que propose Bybee (2001, 2005) repose sur l'hypothèse selon laquelle certains mots et certaines combinaisons de mots, les constructions, sont mémorisées dans le lexique. Cette mémorisation ne peut se faire que par la fréquence d'usage (cf. Bybee 1995, 1998, 2005) qui permet de renforcer leur «*force lexicale*» (Bybee, 2005 : 25). Les éléments constitutifs de ces constructions obtenues entretiennent donc entre eux des relations sémantiques particulières, relations inexistantes lorsque ces éléments intègrent d'autres contextes.

Bybee établit que les mécanismes permettant d'établir ces constructions sont au nombre de deux (2005 : 26) :

- (i) automatisation par répétition de « tronçons » de matériau linguistique ;
- (ii) et catégorisation des items qui apparaissent dans certaines positions des « tronçons » plus grands.

Selon Bybee (2005), la liaison est l'un des phénomènes dans lequel s'observe le mieux le lien entre la syntaxe et la phonologie d'une part, et le stockage lexical d'autre part.

Elle propose ainsi que les contextes d'apparition de la liaison soient stockés dans la mémoire. A l'instar des autres travaux, son approche repose sur « *l'intuition* » selon laquelle la liaison est plus souvent réalisée lorsque la cohésion syntaxique est forte. Elle ajoute que cette cohésion est le résultat direct de la fréquence de co-occurrences : « *les mots qui sont souvent employés ensemble semblent liés par une plus forte cohésion* » (Bybee, 2005 : 24).

Cette « *intuition* » s'appuie en partie sur le fait que lorsqu'un contexte de liaison tend à ne plus se réaliser, les séquences de mots les plus fréquentes maintiennent tout de même cette liaison (cf. Ågren, 1973 et Delattre, 1966). Bybee illustre cette tendance en opposant deux types de constructions : d'une part, les constructions de type Déterminant + Nom (par ex. *son avis*), très fréquentes et pour lesquelles la liaison est catégorique, et d'autre part, les constructions de type Nom pluriel + Adjectif (par ex. *théories absurdes*), peu fréquentes du fait qu'il s'agit de « *catégorie lexicale restreinte* » (Bybee, 2005 : 25) et pour lesquelles, la liaison est variable.

Comme le rappelle Dugua (2006 : 48) :

La notion de fréquence tient donc une place centrale dans la théorie de Bybee et se retrouve dans sa conception de l'organisation du lexique. Elle suggère que le lexique est formé d'unités de tailles variables, à savoir des mots ou des suites de mots (elle parle de constructions pour les unités complexes) et que cette constitution du lexique est justifiée par des effets de fréquence. Par conséquent, des suites de mots souvent employées ensemble seront stockées telles quelles dans le lexique, sous la forme d'une construction. La CL n'appartient donc ni au Mot1, ni au Mot2 mais à la construction elle-même, la construction correspondant, pour Bybee, à l'unité de stockage dans le lexique.

Une analyse de la liaison dans un tel cadre permet, toujours selon Bybee, de prédire les contextes de liaison pérennes. Elle considère que, même si la liaison est conditionnée par des motivations phonétiques, les consonnes de liaison, qui n'apparaissent jamais « *indépendamment d'une construction particulière* » (Bybee, 2005 : 25), sont (i) très proches des alternances conditionnées morphologiquement et lexicalement, et (ii) peuvent devenir très productives, se stabiliser ou disparaître selon la fréquence des constructions en jeu (cf. Bybee, 2005) :

Si une construction a une fréquence de type élevé, sa productivité augmente. Les alternances non productives se stabilisent ou disparaissent progressivement, les formes peu fréquentes d'abord, puis celles qui ont

une fréquence d'occurrence élevée et qui résistent le plus longtemps.
(Bybee, 2005 : 30)

De la même façon, Laks (2005b : 117) propose de traiter des liaisons selon leur fréquence positionnelle et de définir les liaisons en fonction de la fréquence de cooccurrence des mots.

Selon lui, la liaison en français « *ne correspond pas [...] à un phénomène mais à deux, de niveaux linguistiques différents* » (Laks, 2005b : 116-117) :

1. les liaisons variables ou plutôt les liaisons « *postposées à la catégorie principale* » qui relèvent d'un traitement phonologique puisqu'elles sont « *liées à l'organisation prosodique, syllabique et rythmique des cursus* » ;

2. et les liaisons catégoriques ou plutôt les liaisons « *préposées à la catégorie principale* » qui relèvent d'un traitement morphologique puisque la liaison est motivée par leur « *fonctionnalité* ».

Les liaisons préposées à une catégorie principale sont vues comme des « *constructions synthétiques stockées comme telles* ». Le possible « *stockage* » de ces constructions est déterminé par l'influence combinée de la fréquence d'usage de la construction et de la fréquence positionnelle. Laks souligne que cette hypothèse a déjà été proposée par Clédat (1917 : 135-175) et qu'elle se vérifie avec les résultats de différentes analyses (par ex. Laks, 1980 ; Durand, Laks et Lyche, 2001 et 2003a ; Lyche, 2004) :

Les liaisons dites obligatoires sont plus fréquentes en style informel qu'en style formel. Un tel effet, totalement inexplicable dans un cadre qui traite la variation comme le produit d'une simple latitude d'interprétation de structures formelles identiques et stables pour tous les locuteurs (Encrevé, 1988), s'explique au contraire aisément par l'interaction des fréquences d'usage et des fréquences positionnelles. (Laks, 2005b : 116)

Il en conclut qu'à la contrainte d'ordre syntaxique est liée une contrainte de fréquence d'usage. Plus les mots sont « *communs* » et « *familiers* » (Laks, 2005b : 116), plus (i) ils sont employés souvent (surtout en style informel) et plus, (ii) ils ont tendance à entraîner la réalisation de la liaison.

A contrario, dans le vocabulaire « *savant* » ou « *rare* », les liaisons se manifestent moins, puisque ces mots sont moins fréquents (utilisation restreinte au style formel) :

Conformément à la loi de Zipf, dans l'usage le moins formel, les constructions sont à la fois moins nombreuses en type, plus fréquentes en nombre et plus spécifiques. C'est pourquoi elles sont constructionnellement plus stables. (Laks, 2005b : 116)

Pourtant, les fréquences d'usage / positionnelles ne nous semblent pas pouvoir expliquer aussi « aisément » les résultats obtenus. Dans un article consacré à l'étude des réalisations des consonnes de liaisons /z/ et /t/ présentes dans une enquête PFC effectuée en Vendée (Mallet, 2003), nous avons effectivement noté que ces deux consonnes de liaison, /z/ et /t/, lient plus dans la conversation informelle que dans la conversation formelle. Cependant, l'écart de réalisation est relativement faible. Par ailleurs, si les taux de réalisation des liaisons catégoriques sont moindres que ceux des liaisons variables, c'est parce que (i) comme le veut la tradition grammaticale, les constructions de type Présentatif + (« *c'est* + ») sont comptabilisées dans les catégoriques, alors que (ii) les résultats de différentes études indiquent que ce contexte est variable (cf. 3.2.3.2.2) :

L'écart de réalisation est toutefois relativement faible, surtout si l'on tient compte du fait que certaines liaisons, classées dans les obligatoires, sont de moins en moins catégoriques et tendent à devenir variables. Ainsi, les présentatifs (*c'est*+) intégreraient la catégorie des liaisons variables, tout comme n'importe quelle forme verbale. (Mallet, 2003 : 149)

Si les présentatifs intègrent le paradigme des variables, il devient alors très compliqué de continuer à :

1. opposer « *liaisons préposées à la catégorie principale* » et « *liaisons postposées à la catégorie principale* » comme synonymes de « *liaisons catégoriques* » vs. « *liaisons variables* » ;
2. et à maintenir que les liaisons catégoriques sont plus nombreuses en style « informel » qu'en style « formel ».

Concernant l'approche de Bybee, la plupart des critiques sont apportées par les acquisitionnistes. Wauquier-Gravelines (2005 : 73-74), par exemple, considère que Bybee confond fréquence lexicale avec contraintes morphosyntaxiques. Par ailleurs, selon cette auteure, les modèles exemplaristes ne peuvent rendre compte de l'acquisition de la liaison. Ramener l'usage à la fréquence d'occurrence pure des objets constitue plus

une démarche quantitative que qualitative ou structurelle et nécessite une modélisation stricte :

L'usage est envisagé par Bybee exclusivement de manière quantitative et jamais qualitative, ni pragmatique, ni structurelle. Il est ramené à la fréquence d'occurrence pure des objets physiques dans le signal (« *token frequency* ») et des structures possibles (par exemple le morphème dérivationnel *-ed* pour le passé en anglais : « *type-frequency* ») que le sujet va percevoir et utiliser en production. Ce qui suppose qu'il y a une corrélation directe entre la forme de l'objet (le son « *ed* ») ou de la structure dans laquelle il apparaît (la dérivation du passé) et le nombre de fois où il a donné lieu à une expérience sensorielle de perception ou de production, indépendamment de ses conditions de réalisation, de ses caractéristiques structurelles intrinsèques et des relations qu'il entretient avec d'autres objets de même nature. [...] La logique du modèle proposé par Bybee nécessite que ces aspects soient modélisés aussi et n'exclut aucunement qu'ils ne puissent jouer aucun rôle dans la constitution des représentations. La logique de ce modèle nous montre bien comment cette conception de la phonologie - et de la grammaire - repose sur un empirisme excessivement réducteur qui simplifie à l'extrême le rapport entre la phénoménologie et la compréhension qu'on peut en avoir, puisque l'usage est ramené à sa pure dimension quantitative.

4.5 Le statut lexical de la consonne de liaison

Comme nous l'avons vu précédemment, les différentes analyses proposées autour de la formalisation de la liaison font émerger un problème central : le statut lexical de la consonne de liaison qui est, de ce fait, l'une des thématiques récurrentes des analyses portant sur la liaison (*cf.* Chevrot et Laks, 2005 : 3).

L'hypothèse la plus couramment soutenue est sans équivoque : (i) le statut lexical de la consonne de liaison ne se pose pas, la consonne est un segment lexicalement rattaché au mot de gauche⁷⁹ (*cf.* Encrevé, 1988 ; Tranel, 1990, 1992 et 1995 ; Paradis et El Fenne, 1995). D'autres positions sont également envisagées : (ii) celle qui rattache la

⁷⁹ Désormais noté M1.

CL au mot de droite⁸⁰ (cf. Morin et Kaye, 1982) ; (iii) celle qui accorde son indépendance lexicale à la CL ; elle est alors considérée comme épenthétique (cf. Tranel, 1981) ; (iv) celle qui considère que les trois options sont possibles (cf. Côté 2005) ; et enfin (v) celle qui base ses analyses dans un cadre de Grammaires de Construction (cf. Bybee 2001 ; Laks 2005b). Quelle que soit l'hypothèse retenue, la consonne de liaison apparaît toujours comme une consonne nécessitant un traitement particulier.

Dans cette section, nous nous attacherons donc à faire le point sur les différentes propositions relatives au statut lexical de la liaison. Du fait qu'elles ont été en partie décrites dans la section précédente, certaines de ces propositions seront résumées ; d'autres seront exposées plus longuement :

1. la CL est rattachée à M1 (cf. 4.5.1) ;
2. la CL est rattachée à M2 (cf. 4.5.2) ;
3. la CL est lexicalement indépendante (cf. 4.5.3) ;
4. enfin, la CL appartient à la construction (cf. 4.5.4).

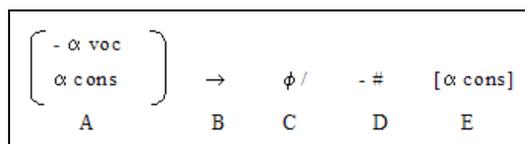
4.5.1 La CL est rattachée à M1

La première proposition accorde à la consonne de liaison une place lexicale, et rattache cette consonne au mot précédent (M1). Trois analyses ont été proposées : la troncation, la supplétion et la flottance.

1. La troncation

Le principe de la troncation est le suivant : la consonne finale d'un morphème peut être effacée soit en finale, soit devant une consonne par une règle générale (de troncation donc). Comme nous l'avons vu en 4.1.1, cette proposition a vu le jour avec les travaux de Schane (1968) qui a proposé la *French Truncation Rule* :

⁸⁰ Désormais noté M2.



2. La supplétion

Dans le cadre d'une analyse supplétive, M1 est considéré comme allomorphique. Il a donc différentes représentations lexicales possibles : (i) soit une forme de liaison (par ex. /les/ dans *les années*) ; (ii) soit une forme de non-liaison (par ex. /le/ dans *les formes*).

La supplétion a été envisagée dans de nombreux cadres théoriques différents : générativiste multilinéaire (cf. Klausenburger, 1984 ; Morin, 1987 ; Green et Hintze, 1988 ; Perlmutter, 1998 ; Tranel, 1995b) ; Théorie de l'Optimalité (cf. Tranel, 2000), etc.

La plupart des travaux se sont plus particulièrement centrés sur l'allomorphie des adjectifs. Pour Tranel (2000 : 50), par exemple, la supplétion se justifie dans le cas des consonnes intrinsèques facultatives, mais pas dans le cas des consonnes latentes.

Selon Côté (2005 : 74), un nombre restreint d'adjectifs doit être analysé par supplétion : les adjectifs présentant une alternance vocalique ([ɔ̃] – [ɔn] ; [ɛ] – [ɛn] et [o] – [ɔ]). Elle suppose que les adjectifs présentent deux formes supplétives : une forme de non-liaison à finale vocalique et une forme de liaison à finale consonantique (par ex. *bon* /bɔ̃/, bɔn/).

L'avantage d'une analyse par supplétion est qu'elle rend compte de la nature de la consonne alternante puisque celle-ci est lexicalisée. En revanche, elle ne tient pas compte du fait que, parfois, la distribution des consonnes de liaison n'est pas aléatoire ; enfin, elle ne permet aucune prédiction quant à la nature de la consonne réalisée.

3. La flottance

Dans une approche autosegmentale (cf. 4.2 et Encrevé, 1988), la consonne de liaison est considérée comme flottante :

Le *segment flottant* correspond à l'analyse selon laquelle la CL appartient lexicalement au mot mais doit avoir une représentation distincte de celle

des consonnes toujours réalisées dans la prononciation : *flottante*, la CL dépend pour sa prononciation de l'intervention, ou non, d'une opération d'ancrage dans le squelette, ainsi que d'une opération d'interprétation syllabique de la position dans laquelle elle s'ancre. [...] Cette représentation lexicale de CL comme *segment flottant auquel correspond un constituant syllabique également flottant ainsi qu'une position du squelette sans association aucune*, exprime bien l'opposition entre notre théorie de la liaison et la théorie « concrète » qui se traduit, elle, par un traitement épenthétique de CL. (Encrevé, 1988 : 170-171)

La CL n'est donc pas ancrée dans le squelette et flotte ; elle n'est associée à aucune position squelettale. Réaliser la liaison revient alors à ancrer cette consonne dans la structure prosodique. Pour ce faire, il est nécessaire que les conditions syllabiques soient remplies (présence d'une position d'attaque disponible à sa gauche). Ne pas réaliser la liaison, ou plutôt ne pas ancrer la consonne dans le site disponible, conserve tout de même à cette dernière son statut de « flottance ».

Considérer la consonne de liaison comme un segment flottant permet de rendre compte de la variabilité de la nature de la CL.

4.5.2 La CL est rattachée à M2

Le traitement de la consonne de liaison comme rattachée au second mot (M2) est de loin la proposition la moins retenue.

En dehors de Ternes (1977) qui propose une analyse supplétive de la consonne de liaison, la plupart des travaux validant le rattachement de la CL à M2 considère cette consonne comme présente au niveau lexical. Ces travaux portent essentiellement sur le /z/ du pluriel (Gougenheim, 1938 ; Morin, 1981 et 1986 ; Morin et Kaye, 1982) et les enclitiques (Morin, 1979a et b, 1986 ; Côté, 2005). On compte également quelques analyses sur les adjectifs prénominaux (Klausenburger, 2001 ; Morin, 2003).

Gougenheim (1938 : 59-60), par exemple, observe que dans une séquence Nom+adjectif, le comportement de la liaison peut se révéler très différent. Comparant *un nez aquilin* (1) et *des nez aquilins* (2), il fait remarquer que la réalisation de la liaison ne peut s'observer que dans le second cas, où elle est alors variable ([denezakilẽ] / [deneakilẽ]). Par contre, dans le premier cas, la liaison est erratique ([ẽneakilẽ] / *[ẽnezakilẽ]). La consonne /z/ se réalise phonétiquement et

morphologiquement comme un préfixe pluriel de l'adjectif (*aquilins*). Cette analyse de la présence de la consonne de liaison à l'initiale de M2 est confirmée par Morin et Kaye (1982 : 320 et *sq.*).

Concernant l'analyse de la liaison avec les enclitiques (par ex. *vas-y*, *va-t-il* et *vont-elles*), Morin (1979a et b, 1986) propose que ces derniers soient lexicalisés avec la consonne qui les précède (*/zi/*, */ti(l)/*, */te(l)/*). La consonne de liaison est analysée comme une consonne initiale de M2 et n'est donc pas à proprement parler une consonne de liaison (*cf.* Côté, 2005 : 73).

4.5.3 La CL est épenthétique

La conception épenthétique considère la consonne de liaison comme indépendante au niveau lexical : elle n'appartient ni à M1, ni à M2 (*cf.* Klausenburger, 1974 ; Morin, 1981 et 1986 ; Wetzels, 1987 ; Côté, 2005).

Klausenburger (1974) postule cette indépendance lexicale de la CL à l'aide d'une règle d'épenthèse (*cf.* 4.1.3, p. 117). Rappelons qu'il rejette la règle de troncation de Schane (1968) et considère la liaison et l'élision comme étant des procédés distincts. Selon lui, la consonne de liaison a un statut lexical différent de celui de la consonne finale fixe.

Côté (2005) fait état du consensus général qui pose que l'encodage de la CL se fait à l'initiale de M2 chez les enfants de 2 / 3 ans⁸¹. Elle relève ensuite que le problème de la détermination du statut lexical de la CL peut apparaître lorsque l'on considère que l'adulte rattache la CL à M1 (*cf.* Morel, 1994 ; Wauquier-Gravelines, 2001). Il devient alors extrêmement compliqué d'expliquer comment l'enfant, lors de son développement, passe d'un encodage sur M2 à un encodage sur M1. D'autres études penchent plus pour un détachement progressif de la CL de M2 jusqu'à l'épenthèse (*cf.* Chevrot et Fayol, 2001 ; Dugua, 2002). Il semblerait que dans les erreurs de liaison des enfants âgés de 2 à 5 ans, le taux d'omission de la liaison catégorique (par ex. : *un ours* réalisé [œurs]) augmente progressivement et qu'il soit positivement corrélé avec le

⁸¹ Cette hypothèse est notamment confirmée par le fait que l'on trouve beaucoup de formes comme */nurs/*, */zarbr/* en début d'énoncé (*cf.* Dugua, 2002).

nombre de liaisons « correctement » réalisées. Par ailleurs, les enfants de 4 ans produisent plus d'erreurs sur les consonnes initiales stables (par ex. *un zèbre* réalisé [œ̃ɛbr]), qu'ils traitent comme des CL.

Pour Côté, seule l'analyse de la CL (chez l'adulte) comme épenthétique permet de rendre compte de ces évolutions (2005 : 69-70). Sa proposition est de considérer que l'omission des CL chez l'enfant est due à la non-application de la règle d'épenthèse. Elle explique le processus de la façon suivante : plus l'enfant progresse dans son acquisition de la liaison, plus le détachement se renforce et plus il omet de liaisons catégoriques. Cette omission est le résultat direct d'une « *surapplication du processus de détachement* » :

L'enfant détache les /n, z, t/ initiaux dans les mots qui possèdent une variante à initiale vocalique mais également au début de quelques mots qui ne présentent pas d'alternances (p. ex. *zèbre*). (Côté, 2005 : 69)

L'avantage d'une analyse par épenthèse est qu'elle permet de rendre compte de l'évolution des erreurs de liaison. Elle est par ailleurs étayée par différentes analyses acoustiques portant sur les CL et par une étude de sujets aphasiques menée par Kilani-Schoch (1983).

Les analyses acoustiques indiquent que les CL ont une représentation spécifique différente des autres consonnes (fixes, finales ou initiales). Par exemple, les analyses de 24 paires de pseudo-mots (par ex. *petit abard / petit tabard*) menées par Dejean de la Bâtie (1993) indiquent que les CL se distinguent des consonnes initiales par leur durée plus courte. Cette différence est également exposée par Spinelli *et al.* (2003). Autrement dit, les CL ne sont rattachées ni à M1, ni à M2 et la solution d'une consonne par épenthèse reste la plus plausible.

4.5.4 La consonne appartient à la construction

Dans le cadre des Grammaires de Construction, Bybee (2001 et 2005) propose que les CL ne soient ni rattachées à M1, ni rattachées à M2 mais sont dépendantes de la construction, autrement dit de « l'unité » stockée dans le lexique.

4.5.5 Bilan

Comme nous venons de le voir dans cette section, les propositions quant au statut de la consonne de liaison sont fortement concurrentes.

Dans un article consacré à la présentation des différentes positions évoquées ci-dessus, Côté (2005)⁸², qui adopte une conception essentiellement épenthétique⁸³, s'interroge sur la difficulté à déterminer précisément le statut lexical de la CL. Selon elle, plusieurs facteurs contribuent à cette « *incertitude* » (2005 : 66) : avant tout, l'intérêt quasi exclusif d'un grand nombre de spécialistes pour les problèmes de formalisation au détriment de l'empirie, mais aussi une pauvreté du matériel empirique examiné :

Le maintien de l'incertitude s'explique en partie par la nature du débat qui s'est largement concentré sur des arguments formels plutôt qu'empiriques, et par l'éventail plutôt limité des données généralement considérées. (Côté, 2005 : 66)

Côté, par le biais de données acquisitionnelles, démontre que l'on peut envisager de donner un statut « mixte » à la consonne de liaison : (i) les consonnes épenthétiques représentent le cas général, (ii) les consonnes initiales de mot pour les combinaisons fréquentes CL+M2 et (iii) les consonnes finales fixes pour les adjectifs. Cette proposition remet en cause plusieurs aspects de la phonologie du français (*cf.* 2005 : 74-75). Elle rejette notamment la notion de consonne flottante et propose que l'alternance des voyelles nasales ([bɔ̃] ~ [bɔ̃n]) dans les adjectifs relève de la supplétion. Les voyelles nasales seraient alors désormais envisagées comme sous-jacentes et non dérivées des formes orales (*cf.* Morin, 1983 ; Tranel, 1990).

⁸² Côté (2005). Le statut lexical des consonnes de liaison. *Langages*, 158, 66-78.

⁸³ Pour Côté (2005) les consonnes de liaison n'ont pas toutes la même place lexicale : majoritairement épenthétiques, un certain nombre d'entre elles peuvent toutefois être rattachées à M2, tandis que d'autres sont en fait des consonnes finales fixes dans des liaisons supplétives.

Seconde partie :

Analyses empiriques

Chapitre 5

Méthodologie

L'objectif de ce chapitre est de présenter la méthode utilisée dans le cadre du projet PFC « Phonologie du français contemporain (PFC) : usages, variétés et structures » (cf. 5.1), et celle que nous avons adoptée pour constituer les données qui seront analysées dans le Chapitre 6. Nous nous intéresserons plus particulièrement au protocole, au codage liaison et aux problèmes récurrents à l'utilisation d'un corpus oral. Dans une seconde section, nous présentons le corpus PFC (cf. pp. 165 et *sq.*).

5.1 Le projet PFC

De nombreux phonologues construisent la base empirique de leurs analyses en puisant dans les descriptions de leurs prédécesseurs pour affiner et enrichir leurs propres données. Peu cependant, dans la tradition générativiste, questionnent la qualité de leurs sources et se préoccupent de savoir si plusieurs variétés de langue coexistent chez un même locuteur ou dans une même région donnée. Face à cette situation qui peut aboutir à des analyses de données erronées (cf. Morin, 1987 pour une discussion sur le sujet), le projet PFC s'est proposé de constituer un corpus de référence à même de nous renseigner sur la prononciation du français saisie dans sa variation et dans la réalité de ses usages attestés. Cette variation, qu'elle soit stylistique, géographique, sociale ou inhérente au système, est omniprésente et son traitement s'inscrit parmi les principaux enjeux de la linguistique moderne.

Le projet PFC est un projet fédératif international placé sous la coordination de Jacques Durand, Bernard Laks et Chantal Lyche, dont l'objectif final est de mettre à

disposition de la communauté scientifique des données documentant la prononciation du français dans sa variation. PFC part du principe qu'il est nécessaire de poursuivre le travail de description entrepris par les sociolinguistes et s'assigne les objectifs suivants :

1. fournir une meilleure image du français parlé dans son unité et sa diversité ;
2. mettre à l'épreuve les modèles phonologiques sur les plans synchroniques et diachroniques ;
3. constituer une base de données sur le français oral suffisamment importante pour permettre la conservation du patrimoine linguistique des espaces francophones du monde ;
4. élargir et surtout renouveler les données et les analyses pour favoriser l'avancement de l'enseignement du français et de la linguistique française ;
5. favoriser les échanges entre les connaissances phonologiques et les outils de traitement de la parole ;
6. et enfin proposer des analyses phonologiques de phénomènes variables.

Etant donné ces objectifs, la question d'un choix méthodologique et théorique entre « linguistique de corpus » et « introspection » ne se pose pas ; PFC ne se situe pas dans cette alternative en ce sens que « théorie » et « corpus » n'ont pas les mêmes rôles à jouer, même s'ils se complètent. La théorie est abstraite, tente de rendre compte du système, du possible de langue. Elle construit des modèles le plus souvent à partir de l'introspection. Le corpus, quant à lui, tente de révéler l'événementiel, la réalité de l'usage. Pour une discussion précise sur le sujet, voir Milner (1989), Labov (2002) et Delais-Roussarie (2008).

PFC vise, à terme, l'élaboration d'un grand corpus oral de plus de 600 heures d'enregistrements comprenant une ample couverture géographique à l'intérieur et à l'extérieur de l'Hexagone (par ex. la Belgique, la Suisse, le Canada, divers pays de l'Afrique francophone, la Louisiane, *etc.*).

Pour atteindre ses objectifs, le corpus PFC a été établi à partir d'un protocole⁸⁴ qui doit permettre l'obtention de données strictement comparables pour un certain nombre

⁸⁴ www.projet-pfc.net ; Durand *et al.*, 2002 ; Durand et Lyche, 2003 ; Durand et Lyche, 2004 et les Bulletins PFC 1-6. Les Bulletins PFC sont disponibles en ligne sur le site du projet.

de phénomènes caractéristiques de la phonologie du français, principalement les inventaires phonémiques, le schwa et la liaison.

5.1.1 Protocole d'enquête

Pour documenter les différentes prononciations du français dans le monde francophone, le projet s'appuie sur la notion de « point d'enquête ». Un point d'enquête est délimité selon une certaine zone géographique (par ex. « Treize Vents »), éventuellement complétée de certaines caractéristiques sociales (par ex. « Paris noblesse »).

Pour chaque point d'enquête, le protocole recommande aux enquêteurs de constituer un groupe d'une douzaine de locuteurs. Un des principaux critères requis dans le « recrutement » de ces locuteurs est qu'ils aient un bon ancrage dans la communauté de référence choisie. Pour cela, l'enquêteur doit choisir des locuteurs qui ont vécu toute leur vie ou presque dans la région où l'enquête est effectuée.

Le groupe constitué doit présenter :

- (i) une répartition égale entre les sexes (moitié femmes / moitié hommes),
- (ii) une représentativité des différentes tranches d'âge (idéalement répartis entre les 18-30 ans, les 30-45 ans, les 45-60 ans et les plus de 60 ans ; si l'on n'atteint pas les 4 tranches d'âge, deux tranches minimum sont requises ;
- (iii) et autant que possible, une représentativité sociale (qui n'est cependant pas exigée par le protocole).

La procédure d'enquête comporte deux volets auxquels tous les locuteurs sélectionnés doivent se soumettre : une conversation de groupe et une lecture de documents. La conversation se scinde en deux parties : une discussion informelle et une entrevue formelle. La lecture de documents comporte deux étapes : la lecture d'un texte et la lecture d'une liste de mots.

Pour résumer, le protocole est concrètement constitué d'une interview d'environ une heure par locuteur et se divise en quatre étapes :

- un entretien guidé (discussion formelle) d'environ vingt minutes ;

- un entretien libre (discussion informelle) d'environ vingt minutes ;
- la lecture d'une liste de mots⁸⁵ ;
- la lecture d'un texte.

Concernant la liaison, PFC tente d'approcher au plus près de trois registres différents : (i) la lecture du texte qui se rapproche du style « conférence » proposé par Delattre (1966), (ii) la conversation libre ou informelle qui correspond au registre « familial » ; et (iii) la conversation guidée qui donne un ton plus formel au dialogue et qui se rapproche donc du registre « soigné » de Delattre ou du registre « soutenu » de la classification traditionnelle (*cf.* 3.3.3.1, pp. 88 et *sq.*). Soit :

	Classification traditionnelle des niveaux de langue	Fouché (1959)	Delattre (1966)	PFC
ORAL / ECRIT	/	/	récitation de vers	lecture de texte
	/	/	conférence	
ORAL	Soutenu	soutenu	soigné	conversation guidée
	Standard	courant	familier	conversation libre
	Familier			
	Populaire	/	/	/

Figure 5-1 Répartition des corpus de PFC en fonction des registres

Afin de répondre aux exigences du protocole, l'enquêteur doit avoir une bonne connaissance du terrain d'enquête et les enquêtés doivent – idéalement – faire partie de son réseau social. Le protocole recommande même de travailler en équipe de deux, l'un des deux membres entretenant des relations personnelles avec les enquêtés.

Fonctionner ainsi permet (i) à l'enquêteur le plus éloigné des informateurs d'effectuer les entrevues formelles et d'établir peu à peu, grâce au second enquêteur, une relation de confiance envers les enquêtés ; et (ii) à l'enquêteur le plus proche de

⁸⁵ Pour le test de lecture (*cf.* Annexe 1), il est nécessaire que les informateurs soient relativement bien scolarisés, ce qui exclut certains enquêtés pour lesquels le français a surtout été transmis oralement.

l'enquêté de réaliser les entrevues plus informelles (Durand, Laks et Lyche, 2002b : 8). Aussi, la confiance établie lors de la conversation formelle permettra que l'épreuve de lecture se déroule dans un certain climat de sécurité. Il est donc également nécessaire que l'enquêteur principal se sente à l'aise dans le milieu où il opère. De plus, jouer avec les liens de connaissance permet d'étudier des réseaux denses sur le plan interpersonnel et d'avoir accès au vernaculaire des sujets⁸⁶, ce qui permet d'atténuer les effets de ce que Labov (1972a) appelle le « paradoxe de l'observateur »⁸⁷.

Pour chaque enquêté, une fiche signalétique (*cf.* Annexe 3) est remplie. Sa fonction est de collecter un maximum d'informations socio-démographiques sur les enquêtés (âge, sexe, situation familiale, parcours scolaire et professionnel, intégration dans la communauté étudiée, compétences linguistiques, *etc.*). Les divers renseignements fournis par cette fiche doivent permettre de comprendre, voire d'expliquer les productions langagières des interviewés. Pour autant, ils ne prétendent pas à l'exhaustivité requise dans un cadre de sociolinguistique variationniste. Par ailleurs, le nombre de locuteurs de chaque point d'enquête est trop restreint.

Chacun des participants au projet signe également un consentement de participation (*cf.* Annexe 4), accordant le droit au projet PFC d'exploiter les données obtenues dans un cadre scientifique et garantissant en même temps l'anonymat de l'interviewé.

Cet anonymat est assuré par un code alphanumérique donné à chaque témoin. Les premiers éléments du code indiquent l'origine géographique du locuteur : numéro du département (ou les initiales du pays et de la ville lorsque l'enquête est hors du territoire français) et une lettre qui précise le point d'enquête lorsque plusieurs études sont réalisées dans le même département. Les trois derniers éléments « identifient » la personne : ses initiales (prénom et nom), suivies d'un chiffre permettant de distinguer les personnes ayant les mêmes initiales (*cf.* Durand et *al.*, 2002 : 72).

Nous présentons ci-dessous une illustration de ce codage :

⁸⁶ A ce propos, voir Milroy (1980 : 47).

⁸⁷ « *Observers' Paradox : the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed ; yet we can only obtain these data by systematic observation.* » (Labov 1972a : 209)

Enquête	Code enquête	Code locuteur
Dijon	21a	21ama1 21ash1
Liège (Belgique)	bla	blajv1 blajp1

Figure 5-2 Exemple de codes des points d'enquêtes et des locuteurs

5.1.1.1 L'entretien guidé

L'entretien guidé, comme nous l'avons indiqué précédemment, est généralement mené par l'enquêteur le moins « proche » du témoin. Il s'agit d'un entretien semi-dirigé dans lequel la conversation doit être ouverte. Le protocole⁸⁸ suggère quelques thèmes de discussion proposés sous forme de questions. Leur but est de susciter un « récit de vie ». L'enquêteur peut, par exemple, demander des détails sur la scolarité de son informateur, ou bien encore, se renseigner sur son travail et les relations qu'il entretient avec ses collègues. Le but est d'éviter les questions fermées du type « *êtes-vous né dans cette région ?* », qui entraînent des non-dialogues.

L'enquêteur doit donc obtenir de l'interviewé un récit de vie, en établissant une relation basée sur la confiance et en le stimulant dans ses réponses. Cet entretien permet aussi de compléter la fiche signalétique de l'enquêté⁸⁹.

5.1.1.2 L'entretien libre

L'entretien libre a pour objectif de toucher au plus près le vernaculaire de l'enquêté. Pour cela, il est essentiel de réussir à faire oublier au témoin la présence du magnétophone. Le plus souvent, la conversation se fait à plus de deux personnes, ce qui permet une interaction plus riche.

La principale difficulté de cette méthode est qu'il peut s'avérer compliquer de « contrôler » la discussion. Veiller à ce qu'aucune autre personne que le témoin

⁸⁸ Cf. Bulletin PFC 1.

⁸⁹ Par le biais de cet entretien guidé, certains enquêteurs recueillent des informations utiles à leur propre travail de recherche. Ainsi, Sobotta (2006 : 118) : « *L'entretien guidé présente un autre avantage pratique : il m'a permis en même temps d'interroger les informateurs sur une gamme de questions sociolinguistiques (leur identité régionale, leurs attitudes vis-à-vis du créole et de l'occitan, leurs représentations sur les accents, etc.)* ».

n'accapare entièrement le temps de parole ou à ce que les participants n'en viennent pas à parler tous en même temps, reste une entreprise délicate et parfois aléatoire. En tout cas, il importe d'y être particulièrement sensible, car les données recueillies dans ce genre de situations peuvent s'avérer inexploitable.

5.1.1.3 La liste de mots

La liste de mots a été conçue dans le but de dresser un premier inventaire phonémique des témoins (oppositions vocaliques, position inaccentuée, statut de <gn>, réalisation de <r>, groupe consonantique <ks>, *etc.*). L'inventaire choisi ne prétend pas à l'exhaustivité mais donne cependant des indications sur l'attitude du témoin par rapport à l'orthographe et à la norme.

Cette liste contient 94 mots (*cf.* Annexe 2), elle est commune à tous les locuteurs. Elle est reproduite avec une police facilement visible et comprend deux parties : 84 mots présentés de façon aléatoire, et 10 mots représentant les paires minimales d'un certain nombre de mots de la première partie. Dans le but d'indexer plus facilement les productions, l'informateur doit également lire le chiffre qui précède le mot. Cette lecture étant fastidieuse, il est recommandé de proposer à l'informateur une interruption de quelques secondes au cours de la lecture.

Tous les termes techniques du type *homéopathie*, *encéphalite* ont été évités. Les mots proposés font partie du vocabulaire courant.

5.1.1.4 Le texte

Tout comme la liste de mots présentée ci-dessus, le texte proposé en lecture (*cf.* Annexe 1) est commun à tous les locuteurs. Ce texte a été construit selon plusieurs critères. Tout d'abord, son contenu ne présente aucune difficulté de compréhension. Il a la forme d'un article de journal court où le style « littéraire » a été soigneusement évité, dans le but de ne pas être ressenti comme trop « régional » par les différents locuteurs qui vivent non seulement dans l'Hexagone, mais également dans d'autres pays francophones (par ex. Belgique, Canada, *etc.*).

Lors de son élaboration, la majorité des segments vocaliques et consonantiques du français ont été inclus, ainsi que tout un ensemble de phénomènes phonologiques typiques du français. Ceci permet de compléter le premier inventaire phonémique du locuteur établi grâce à la liste de mots et de tester des phénomènes du français tels que la liaison et les phénomènes connexes de nasalisation et de « h » aspiré, ou bien encore la prononciation du schwa⁹⁰.

La lecture du texte ne permet pas une étude de la parole spontanée, mais sa composition et la diversité des phénomènes en jeu (suites de schwa, groupes consonantiques complexes, liaisons, glissantes) font que les informateurs sont beaucoup moins conscients de leur performance linguistique que lors de la lecture de la liste de mots. Le style obtenu demeure, bien entendu, soutenu mais cependant moins surveillé⁹¹.

5.1.2 Traitement des enquêtes

5.1.2.1 Principes techniques

En ce qui concerne les modalités d'enregistrement, il est important de noter qu'elles déterminent fortement la qualité du signal sonore. Si les enregistrements ont été effectués dans des chambres sourdes, le signal est généralement d'une très bonne qualité, en comparaison du signal souvent bruité obtenu lorsque les enregistrements se font dans des lieux ouverts « à la volée ». Dans certaines études, comme celles portant sur les modalités d'alignement des cibles mélodiques, un signal bruité est un réel handicap. En revanche, pour travailler sur des phénomènes comme la liaison, la qualité du signal est moins importante, dès lors que les productions langagières demeurent audibles. Aussi, les modalités d'enregistrement du corpus doivent-elles être choisies en fonction de l'objet de recherche. (Delais-Roussarie, 2006 : 5)

Les enquêtes sont enregistrées à l'aide de DAT ou d'enregistreurs minidisques au format WAV mono avec un taux d'échantillonnage de 22KHz. Une fois les enregistrements numérisés et transférés sur ordinateur, on obtient 4 fichiers sonores

⁹⁰ Le protocole fournit d'ailleurs une grille des contextes de liaison ainsi que la liste des contextes de schwa (cf. Bulletin PFC 1)

⁹¹ Par ailleurs, ce texte permet de déterminer le rapport qu'entretient le témoin avec la norme écrite.

(.wav). Chacun de ces fichiers son est associé à un fichier TextGrid. Pour chaque enquêté, on obtient donc 8 fichiers (4 fichiers .wav et 4 fichiers .textgrid).

Les fichiers TextGrid sont des fichiers en texte brut balisé avec des informations concernant le début et la fin des intervalles créés. Un exemple de TextGrid est donné en (Figure 5-3) :

```
intervals [31]:
xmin = 144.0426176256023
xmax = 147.29575454451228
text = "GM: Entre copains, dans un champ, on rigole bien. "
intervals [32]:
xmin = 147.29575454451228
max = 153.6516220528288
text = "GM: On fait cuire euh, on fait cuire un petit goret et
puis euh, on est tous ensemble, on mange tout ça, et puis après
on invite tous les autres, quoi. "
intervals [33]:
xmin = 153.6516220528288
xmax = 157.29072436889754
text = "GM: Moi c'est euh, c'est moi puis mes copains, on fait
ça, tout le temps, tous les ans. "
```

Figure 5-3 Exemple de Textgrid

Les fichiers TextGrid sont créés à partir du logiciel PRAAT⁹², qui permet d'aligner les représentations graphiques sur le signal. Cet outil, gratuit, présente de nombreux avantages tels que transcription, étiquetage et segmentation des données audio, analyses phonétiques et acoustiques au niveau segmental, manipulation et modification du signal de la parole, *etc.*⁹³ Là où PRAAT se révèle particulièrement utile est qu'il permet de créer, pour chaque TextGrid, plusieurs tires. Cette application est utilisée par le projet PFC dans le traitement de ses données.

Le traitement des données PFC dans PRAAT est constitué de deux phases :

⁹² PRAAT est un logiciel développé par Paul Boersma et David Weenink de l'Institut de Phonétique d'Amsterdam. Il est téléchargeable à partir du site <http://www.praat.org> et disponible pour les plateformes les plus courantes (Windows, Mac Os, Linux x86). Cet outil permet l'alignement des représentations graphiques sur le signal. Un tutorial et des informations supplémentaires sont disponibles sur les sites suivants : <http://person.sol.lu.se/SidneyWood/praaate/frames.html> et <http://www.fon.hum.uva.nl/praat/>.

⁹³ Pour une description complète, voir Delais-Roussarie, Meqqori et Tarrier (2003).

1) une transcription orthographique, qui constitue la « couche zéro » (première tire), à partir de laquelle les codages schwa (tire 2) et liaison (tire 3) sont réalisés. Le texte et la liste de mots sont entièrement transcrits. Les conversations libre et guidée sont chacune transcrites sur 5 minutes.

2) un codage schwa et un codage liaison. Chaque tire est en fait une reduplication de la première tire (orthographique) sur laquelle des informations alphanumériques sont ajoutées. Les codages schwa et liaison ne sont effectués que pour le texte lu, la conversation guidée et la conversation libre.

La Figure 5-4 illustre une transcription de conversation dans PRAAT ainsi que les codages associés :

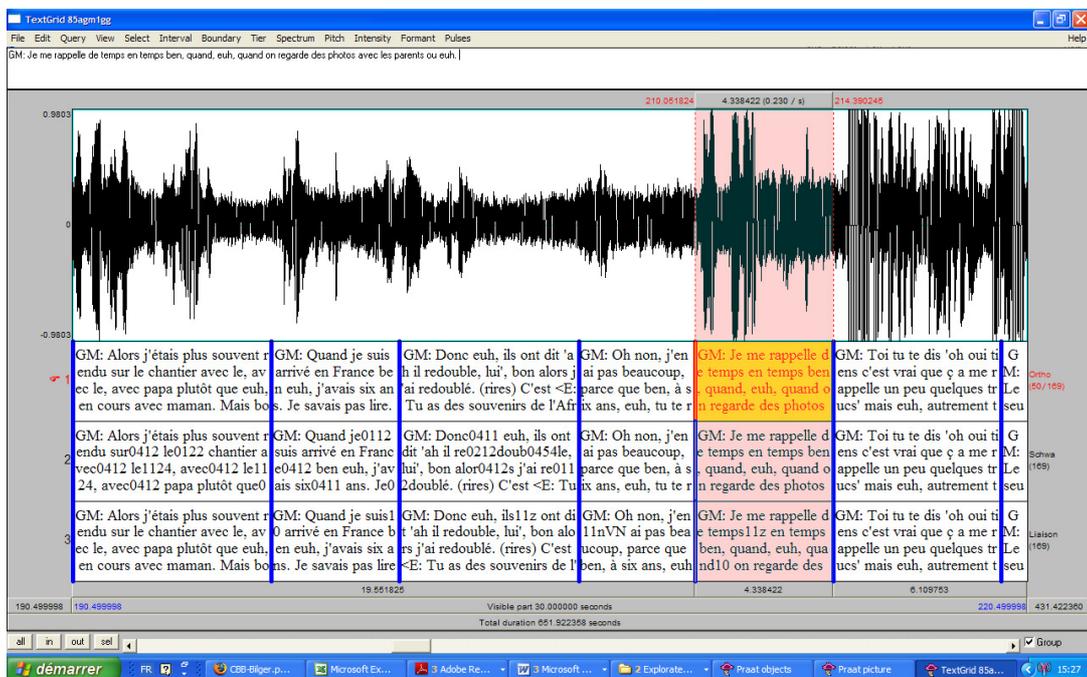


Figure 5-4 Exemple de transcription / codage dans Praat

5.1.2.2 Système de transcription adopté

La transcription des données sonores apparaît « nécessaire » pour que les « choses deviennent manipulables » (Encrevé, 1997 : 110). Il est vrai que dès que l'on s'attache au traitement de données sonores, certaines difficultés apparaissent. Une fois les enregistrements collectés selon un protocole précis, comment opérer le traitement et

analyser les phénomènes qui nous intéressent ? Doit-on étudier les objets uniquement à partir des données sonores ou faut-il faire appel à un système de transcription ? Si le choix se porte sur la transcription, doit-elle être phonétique ou orthographique ? Certains choix explicites doivent se faire. Le projet PFC a opté pour l'utilisation d'une transcription orthographique proche de la norme orthographique. Ce choix, mûrement réfléchi, d'écarter une transcription « phonétique », repose sur plusieurs arguments sur lesquels nous nous proposons de revenir.

D'une part, PFC ne vise pas simplement à fournir une masse conséquente de données orales. Son ambition va plus loin : construire une véritable base de données formatée et annotée qui pourra être utilisée par l'ensemble de la communauté scientifique d'aujourd'hui et de demain. La transcription orthographique présente l'avantage de permettre la mise en circulation réelle d'un corpus, de le rendre accessible et exploitable par d'autres chercheurs :

Cette couche zéro, du fait même qu'elle ne prétend pas refléter fidèlement la structure phonique, permet une exploration beaucoup plus rapide des enregistrements et en retour plus précis sur les données chaque fois que cela s'avère nécessaire. (Durand et Tarrier, 2006 : 142)

Il semble en effet utopique d'envisager une transcription à partir de symboles tels que ceux proposés par l'Alphabet Phonétique International pour un projet à l'échelle de PFC, décentralisé qui plus est. La nécessité d'avoir des données facilement accessibles, comme nous l'avons évoqué précédemment, et la volonté d'un minimum de présupposés théoriques ne peuvent être assurées que par le recours à une transcription orthographique.

D'autre part, Durand et Lyche (2003) rappellent que les transcriptions phonétiques font généralement l'objet de désaccords intersubjectifs quant aux symboles et diacritiques à utiliser et posent de grandes difficultés⁹⁴ :

Si on adopte une transcription large, ou plus précisément phonémique, on met la charrue avant les bœufs : on suppose qu'on a déjà découvert le système qu'on cherche à établir à travers l'enquête. Si on privilégie au contraire des transcriptions étroites (de type allophonique), on se heurte à

⁹⁴ A ce sujet, voir Durand et Tarrier (2006).

une difficulté de taille. Quel degré de finesse phonétique doit-on adopter ? En effet, au niveau phonétique, de nombreuses réalisations ne correspondent pas à des choix binaires mais à des valeurs sur des échelles continues. Ainsi la longueur des voyelles au sein des énoncés n'est pas simplement une opposition entre long (:) et bref, ni même entre long, demi-long et bref, mais entre divers degrés de longueur que seule une étude acoustique précise peut résoudre. De même, l'assimilation de voisement n'est pas toujours un simple changement de valeur (voisé vs. non voisé) mais un phénomène scalaire et qui, selon divers spécialistes, entre en interaction avec d'autres paramètres phonétiques comme la dimension tendu / lâche [...] Ce travail ne peut être fait de manière fiable qu'en utilisant des outils de mesure acoustique. Mais s'engager dans une étude acoustique demande beaucoup de temps et d'énergie même pour un corpus de petite taille et exige de toute façon un déblayage du terrain. (Durand et Lyche, 2003 : 18)

Une transcription entièrement phonétique peut être délicate pour ne pas dire impossible à manipuler par certains linguistes qui travaillent par exemple sur la morphologie, la syntaxe, *etc.* La transcription orthographique de corpus selon des conventions explicites rend leur échange plus facile. Et, bien que la transcription orthographique soit souvent insuffisante pour l'analyse des données, elle donne au chercheur une grande latitude quant au phénomène qu'il compte étudier.

Outre ces deux arguments, rappelons enfin qu'une transcription orthographique standard présente l'avantage « *d'indexer rapidement toutes les occurrences d'une unité lexicale ou de séquences de mots données pour en examiner la prononciation* » (Durand et Tarrier, 2006 : 146).

5.1.2.3 Conventions de transcription

Les conventions de transcription adoptées par PFC s'inspirent des travaux menés autour du corpus du GARS⁹⁵ et de VALIBEL⁹⁶ et correspondent à une transcription orthographique standard qui se veut la plus proche possible des conventions de l'orthographe ordinaire. Les conventions sont strictes et permettent d'assurer l'homogénéité du corpus⁹⁷. Le format de rendu de la transcription sous PRAAT (fichier

⁹⁵ Voir le site <http://www.up.univ-mrs.fr/delic/corpus/index.html>. Aix (Université de Provence).

⁹⁶ Voir le site <http://valibel.fltr.ucl.ac.be/>. UCL (Louvain-la-Neuve).

⁹⁷ Le respect des conventions par tous les participants est d'autant plus important que PFC est un projet décentralisé.

TextGrid aligné sur le signal) peut être rendu plus « lisible » grâce à sa conversion en fichier .txt par l'utilitaire TRANSPRAAT⁹⁸. Le nouveau fichier créé (Figure 5-5) n'est alors plus aligné sur le signal sonore et prend la forme suivante :

E: Et qu'est-ce que vous faites alors de vos week-ends, de vos soirées ? GM: Alors nos soirées, boîtes de nuit (rires). Boîtes de nuit, et puis euh, concerts de temps en temps. Non autrement ben euh, oui souvent, souvent boîtes. Un peu concerts, mais de moins en moins. Autrement ben on fait des fêtes entre nous, tous les étés on fait un méchoui. Entre copains, dans un champ, on rigole bien. On fait cuire euh, on fait cuire un petit goret et puis euh, on est tous ensemble, on mange tout ça, et puis après on invite tous les autres, quoi.
--

Figure 5-5 Exemple de transcription « transpraatée »

Les conventions de transcription sont partiellement présentées ci-dessous⁹⁹ :

1. Formes de mot

Au niveau lexical, les différentes réalisations d'un mot sont ramenées à la forme standard de ce mot en contexte. Par ex. : [ptit], [ptitə], [pœtit] ou [pt.ɪt] seront transcrits « *petite*' ; [ʒə sɥi], [ʒsɥi], [ʃsɥi] ou [ʃɥi] *etc.*, seront transcrits « *je suis*'. En revanche, les éléments lexicaux absents ne sont pas réintroduits. Par ex. [ivwapasakɔmsa] sera transcrit « *il voit pas ça comme ça* » et non « *il ne voit pas ça comme ça* »¹⁰⁰.

2. Tours de parole

L'enquêteur est désigné par E ou E1, E2 *etc.*, s'il y a plusieurs enquêteurs. Les témoins sont identifiés par leurs initiales (par ex. : GM). Les initiales sont suivies de deux points (:) et sont reprises à chaque intervalle même s'il n'y a pas de changement de locuteur. Il n'y a aucun paragraphe, ni retour à la ligne.

3. Ponctuation

Une pause brève (mouvement mélodique indiquant une continuation) est indiquée par une virgule, une pause plus longue (fin d'énoncé indiquée mélodiquement) par un

⁹⁸ TRANSPRAAT est un logiciel développé par Meqqori (ERSS, Toulouse Le Mirail) et disponible sur le site PFC : www.projet-pfc.net.

⁹⁹ Pour une description complète des conventions adoptées par le projet, cf. Bulletin PFC 1.

¹⁰⁰ Le protocole recommande de mettre quand même le *ne* quand on ne peut pas être sûr auditivement de sa présence/absence. Par ex. *on n'atteint pas l'ivresse*.

point, une question par un point d'interrogation. Il n'y a ni point d'exclamation, ni point-virgule.

4. Chevauchements

La parole du locuteur qui interrompt le discours de l'autre est transcrite entre chevrons (< >) (hormis les remarques discursives du type *oui, hum, pff, etc.* qui sont ignorées et non transcrites). Par ex. : *AB : Tu as vu Pierre ces derniers <E : Pierre, non.> temps ?*

5. Troncations

Les troncations sont indiquées par une barre oblique suivie d'un espace. Par ex. : *Il m'a pro/ promis de revenir vite.*

6. Répétitions

Les répétitions sont séparées par une virgule. Par ex. : *On, on fait ça tout le temps.*

7. Sigles

Les sigles épelés sont transcrits avec chaque lettre séparée par un point (par ex. : *la S.N.CF. □laesenseef□ est encore en grève*) et les sigles lus sont transcrits comme un mot (par ex. : *Le CNET [snet] embauche*).

8. Hésitations

Les hésitations sont transcrites *euh*.

9. Discours rapporté

Le discours rapporté est signalé par des guillemets simples (‘ ’) en début et en fin de discours. Par ex. : *Il m'a dit ‘c'est comme une prépa, nuit et jour non stop et basta’, alors je l'ai envoyé sur les roses.*

10. Parenthèses

Les syllabes incompréhensibles sont indiquées par des X (maximum 3) entre parenthèses. Par ex. : *Paul s'en va et juste à ce moment, (XXX) et il tombe sur lui.* Les parenthèses sont également utilisées pour des commentaires. Par ex. : *Paul s'en va et juste à ce moment (bruit) (catastrophe) et il tombe sur lui.*

5.1.2.4 La transcription orthographique et ses limites

Bien que la transcription orthographique standard constitue le cadre idéal pour l'entreprise de PFC, les participants au projet n'en sont pas moins conscients qu'un certain nombre de problèmes peuvent émerger :

Aussi réels que soient les possibilités et avantages d'une TOS [Transcription Orthographique Standard] alignée sur le signal de la parole, il serait erroné de croire que celle-ci puisse répondre à toutes les situations que peut rencontrer le transcripateur. Il existe en effet différentes circonstances où une TOS ne permet pas de satisfaire pleinement aux critères de non-altération et de récupérabilité de l'information. [...] Ces situations [...] ont quasiment toutes en commun de comporter des réalisations dont la transcription nécessite cruciallement une référence explicite à l'oralité et à ce qu'on a pu appeler sa nature polyphonique. (Durand et TARRIER, 2006 : 147)

A titre d'illustration, prenons par exemple le problème de l'interaction des langues, soit « l'irruption de formes linguistiques étrangères ». Les enquêtes PFC l'attestent. Lorsque la langue utilisée possède elle aussi une orthographe standardisée dont la typographie est équivalente à celle du français (par ex. le basque), le transcripateur peut utiliser l'orthographe de la langue en question. Par contre, pour l'ensemble des langues utilisant un système d'écriture différent (par ex. l'arabe), PFC préconise le recours à l'utilisation d'un alphabet phonétique (API ou SAMPA).

Le recensement de ces difficultés n'a pu être opéré qu'au fur et à mesure de la constitution du corpus PFC. Les solutions envisagées, bien qu'elles puissent sembler partielles, font continuellement l'objet de recherches approfondies. Nous ne développons pas ici l'ensemble des problèmes existant, et renvoyons le lecteur à l'article de Durand et TARRIER (2006 : 139-158) pour le détail.

Dans les faits, une enquête sociolinguistique au sens propre n'est pas possible. Les données PFC actuelles, que nous présenterons dans la section 5.2 (pp. 166 et *sq.*), montrent qu'il n'est pas toujours possible d'atteindre cette distribution. Certaines enquêtes n'ont pas le nombre de locuteurs requis, d'autres n'illustrent pas complètement les diversités sociale ou générationnelle voulues, et d'autres encore ont des qualités d'enregistrement qui ne permettent pas forcément leur exploitation.

5.1.3 Le codage liaison

Étant donné notre sujet d'étude, nous n'aborderons dans cette section que la question du codage de la liaison. Nous ne détaillerons pas les principes du codage schwa. Pour une description complète, voir Durand, Laks et Lyche (2002b : 7-19) ; Durand et Lyche (2003). On se reportera également aux différents bulletins PFC et aux différents travaux de thèse existants sur la question (par ex. Turcsan, 2005 : 48-52 ; Eychenne, 2006 : 38-45).

5.1.3.1 Principes généraux

L'enquête PFC fournit des résultats d'une grande généralité sur la présence ou l'absence de la liaison. Les données recueillies et codées sont de deux ordres : celles obtenues grâce au texte lu et celles de la conversation (conversations guidée et libre). De la même façon que pour le codage schwa, le codage liaison est appliqué à la transcription orthographique dans une tire spécifique intitulée *codage-liaison* (tire 3, cf. Figure 5-4 : 154 pour une illustration). Il est effectué sur le texte lu, sur 5 minutes de transcription de la conversation libre et sur 5 minutes de la transcription de la conversation guidée.

Ce codage constitue une étape préliminaire à toute analyse. Il permet de parcourir l'ensemble du corpus et de disposer de données statistiques de toutes les liaisons observées dans l'analyse auditive d'un enregistrement. Une étude quantitative par extraction des données devient alors possible. Le classement de ces données doit permettre par la suite des analyses plus précises du phénomène :

De tels codages ne constituent qu'une étape préliminaire de toute analyse mais cette étape nous semble indispensable. Il s'agit de balayer largement le corpus et de recueillir des données statistiques qui permettront quelques grandes généralisations et qui serviront d'assise à des analyses plus fines. Une fois le corpus constitué, nous serons en mesure de comparer les prononciations d'un très grand nombre de locuteurs sur des bases similaires. La mise au point d'utilitaires informatiques dédiés à des tâches de comptage et de classement de ces phénomènes permet d'automatiser et d'uniformiser les observations. Les variétés de français ont rarement été abordées avec des principes de codage aussi explicites que les nôtres et nous pouvons donc espérer obtenir des données d'envergure comparables. (Durand et Lyche, 2003: 34)

Le codage est alphanumérique et s'effectue à la suite des lexèmes / graphèmes. Il a été élaboré à partir du fascicule de Pierre Delattre (1951), destiné à un public d'apprenants anglophones, dans lequel un tableau détaillé liste les contextes de liaisons classées comme obligatoires, facultatives et interdites.

Dans le codage liaison, toute consonne graphique finale susceptible d'être une consonne de liaison est codée. Au sens de PFC, par « consonne de liaison », il faut entendre :

[...] toute consonne latente prononcée lorsque le mot suivant a une initiale vocalique : *mes amis* (*mes [z]amis*), *petit ami* (*petit [t]ami*), *toujours amis* (*toujours [z]amis*). Il n'y a pas liaison dans *avec elle* (*avec [k]elle*) mais enchaînement tout comme dans *toujours amis* prononcé *toujours [r]amis*. La liaison peut être enchaînée ou non. Elle est enchaînée lorsque la consonne apparaît dans l'attaque de la première syllabe du mot suivant et non enchaînée lorsqu'elle est prononcée en fin de dernière syllabe du mot qui déclenche la liaison. [...] Sera donc codée toute consonne graphique finale susceptible d'être une consonne de liaison. Une consonne graphique n'est cependant pas toujours à l'origine d'une liaison et il y a aussi liaison dans *quatre enfants* (*quat' [z] enfants*), *il va à Paris* (*i' va [t]à Paris*). Ces cas de liaisons seront aussi notés tout comme les cas (rares) de liaison devant initiale consonantique *on peut venir* (*on peut[t] venir*). (Durand, Laks et Lyche, 2001: 72)

Par ailleurs, le protocole exclut du codage deux contextes décrits par Delattre comme interdits et où la liaison n'est apparemment jamais observée :

- **substantif singulier + « X' »**. Par ex. : *un chamelier émirati, un chien errant.*
- **conjonction *et* + « X' »**. Par ex. : *un chameau et un âne du Poitou.*

Cependant, ces liaisons seront codées si elles sont attestées.

5.1.3.2 Codage alphanumérique

Les conventions du codage liaison sont énumérées ci-dessous. Tous les codages comporteront au minimum deux symboles :

Le premier chiffre indique la longueur (nombre de syllabes) du mot liaisonnant :

- 1 = une syllabe
- 2 = deux syllabes ou plus

Le second chiffre est tiré de l'ensemble {0, 1, 2, 3, 4} où :

- 0 = absence de liaison
- 1 = liaison enchaînée
- 2 = liaison non enchaînée
- 3 = incertitude
- 4 = liaison épenthétique

Ces deux chiffres peuvent être complétés :

1. par la lettre symbolisant la nature de la consonne de liaison (*z, n, t, etc.*) ;
2. par le symbole *h* qui indique l'existence d'une pause-hésitation et/ou d'un coup de glotte ;
3. par les symboles VO (voyelle orale) ou VN (voyelle nasale) qui indiquent la nasalité de la voyelle précédant la CL /n/ réalisée.

Lorsque la liaison est réalisée devant une consonne, le codage prendra la forme de *lxC* où *l* indique qu'une liaison se manifeste, *x* indique la consonne réalisée (*z, t, n, p, etc.*) et *C* indique que cette liaison se produit devant une consonne :

(27) quand c'est les miennes : signal auditif = [kãtselemjen] → codage = <quand11tC c'est les miennes> où la consonne finale de quand est réalisée [t].

Nous proposons en Figure 5-6 un exemple concret de codage liaison qui correspond à un extrait codé de la conversation guidée d'un locuteur de la Vendée (gm1). Les exemples (28), (29) et (30) donnent l'interprétation des trois codages (en gras, soulignés) que nous avons choisi de décrire :

GM: Avec mes copains, on rigole bien. On boit¹⁰ un petit peu, mais pas beaucoup (rires). Autrement ben on fait des fêtes^{10h} entre nous, tous les^{11z} étés²⁰ on fait¹⁰ un méchoui. Entre copains, dans^{11z} un champ, on rigole bien. On fait cuire euh, on fait cuire un petit goret et puis euh, on¹ lⁿ VN est tous ensemble, on mange tout ça, et puis¹⁰ après, on invite tous les^{11z} autres, quoi.

Figure 5-6 Exemple de codage liaison

(28) *les étés on* : signal auditif = [lezeteõ] → codage = <étés²⁰ on> où étés²⁰ est polysyllabique (2) et où la liaison n'est pas réalisée (0)

(29) fêtes entre nous : signal auditif = [fetẽtr̃nu] → codage = <fêtes10h entre>
où fêtes10 est monosyllabique (1) du fait de la chute de schwa, où la liaison n'est pas
réalisée (0) et où il y a un coup de glotte (h)

(30) les étés : signal auditif = [lezete] → codage = <les11z étés> où les11z est
monosyllabique (1) et où la liaison est réalisée (1) en [z]

5.1.3.3 Les limites du codage

Dans cette section, nous nous concentrons sur les problèmes spécifiques que posent la liaison et son codage dans PFC. Ceux-ci nous permettront de discuter des limites du codage de manière plus générale, en montrant que les choix méthodologiques restent pour une part dans la dépendance des choix théoriques.

Tout système de codage comporte en lui-même des limites, car il repose sur le choix de coder ou non tel ou tel aspect de telle ou telle manière. La question est alors d'évaluer les aspects codés, les aspects non codés mais aussi la qualité de la réalisation du codage. C'est souvent « après coup » que l'on se rend compte des problèmes. PFC n'échappe pas à « la règle ». On peut en effet identifier cinq principaux problèmes.

5.1.3.3.1 Le transcripneur codeur qui fait des erreurs

Le transcripneur codeur peut oublier de coder des contextes (avec liaisons variable ou catégorique) ; il peut se tromper de code ; il peut coder des liaisons erratiques non réalisées, *etc.* Il est bien évident qu'on ne peut en dresser la liste *a priori*. On « tombe » sur ces différents problèmes dès que l'on s'attache très précisément aux données. Si une évaluation quantitative générale ne peut être établie, pour les données concernant la liaison, il nous semble toutefois que leur nombre reste marginal.

5.1.3.3.2 Le manque d'informations

Le codage proposé fournit une information minimale. Il joue au mieux le rôle d'un baliseur ce qui, par ailleurs, est cohérent avec les objectifs du projet PFC qui ne vise qu'à fournir (i) les contextes de liaison sous forme de données quantitatives qui pourront être utilisées dans le cadre d'une analyse plus fine et (ii) un codage avec le minimum de présupposés théoriques.

On pourrait regretter l'absence de codage faisant appel à des informations syntaxiques. En effet, il y a un consensus sur le fait que la nature syntaxique des éléments qui peuvent lier est l'un des facteurs principaux qui déterminent la réalisation de la liaison. Rappelons que pour Delattre (1955 : 45), « *la liaison facultative est en partie gouvernée par le degré d'étroitesse de l'union entre les mots* ». De la même façon, Dell (1973) présente les réalisations de liaison en français comme dépendantes de deux conditions qui doivent être remplies simultanément. D'une part, il y a une condition d'ordre phonologique qui stipule que le mot à consonne latente doit être suivi d'un mot commençant par une voyelle. D'autre part, il existe une condition syntaxique qui stipule que les deux mots doivent être étroitement liés sur le plan syntaxique.

Or, le codage mis en place dans PFC ne fait aucune place aux informations de nature syntaxique. Mais si l'on se rappelle que les informations codées doivent être consensuelles et faire appel à un minimum de théorie, on comprend alors pourquoi des informations de ce type n'ont pas été codées. Elles sont laissées au choix des chercheurs susceptibles de travailler en détail cette question.

5.1.3.3.3 Le problème de la granularité du codage

Le problème de la granularité peut être illustré par le codage de la longueur des mots. Il peut apparaître que l'information soit trop succincte. Si le codage fait la différence entre mots monosyllabiques et mots polysyllabiques, il traite de la même manière tous les mots polysyllabiques, quelle que soit leur longueur.

5.1.3.3.4 Le problème de la complétude des contextes

Tous les contextes ne sont pas codés. Ainsi, seules les liaisons « catégoriques » et « variables », qu'elles soient réalisées ou non, sont codées, alors que les liaisons erratiques le sont uniquement quand elles sont réalisées. L'ensemble des contextes phonétiques potentiels n'étant pas codés, on ne sait donc pas évaluer avec précision le nombre de sites non réalisés.

Pour illustrer nos propos, prenons l'exemple du hiatus. Si l'on cherchait à établir la part d'influence du hiatus dans la réalisation de la liaison, les codages permettraient de se fonder de façon certaine sur les contextes de liaison catégorique et variable (ces contextes étant toujours codés que la liaison soit réalisée ou non) et sur les contextes

réalisés de liaison erratique. Par contre, on ne pourrait rendre compte des hiatus dans les contextes non réalisés de liaison erratique, du fait même qu'ils ne sont pas codés.

5.1.3.3.5 Les « choix théoriques implicites » opérés par le codeur

Ce point concerne essentiellement les interactions possibles entre le codage liaison et le codage schwa. Dans certains contextes de liaison, avec un mot (possiblement) liaisonnant susceptible de deux prononciations (par ex. : *plus, moins, fait, vingt, huit, etc.*), il est difficile de déterminer si ce dernier comporte une consonne finale fixe ou une consonne finale flottante. Dans le premier cas, la liaison ne sera pas codée, en revanche, on indiquera la présence ou non d'un schwa. Et dans le second cas, la liaison sera codée au détriment du schwa. Le choix de tel ou tel codage est donc dépendant du savoir du codeur et de ses propres interrogations sur la langue. Il faut toutefois noter que ce problème apparaît lui aussi de manière marginale.

5.1.3.3.6 Bilan : des choix méthodologiques dans la dépendance des choix théoriques

PFC considère que ses choix ne constituent ni un positionnement théorique fort, ni un début d'analyse. Les exemples de problèmes rencontrés dans le codage permettent pourtant de se rendre compte que tel n'est pas le cas. Le phénomène n'est en effet pas considéré dans sa généralité et on en a donc une définition hétérogène puisque l'on se situe à la fois dans la norme et dans la théorie :

- i. on ne se limite pas à un contexte phonétique précis, on écarte du codage des contextes considérés comme « jamais réalisés » (*cf.* « Le problème de la complétude des contextes »). Cela signifie donc que l'objet est déjà identifié, que ses caractéristiques sont déjà connues et reconnues ;
- ii. on prend le parti de coder un certain nombre d'informations au détriment d'autres (*cf.* « Le manque d'informations ») ;
- iii. on code partiellement un aspect (*cf.* « Le problème de la granularité du codage »).

Ces choix méthodologiques soulèvent le problème de la collecte de données et de la conception d'un codage « en général ». En effet, la collecte de données présuppose une

sélection raisonnée, donc une théorie. Saussure (1916 : 149) souligne qu'en linguistique, « rien n'est a priori donné, tout est à construire ». Construire un système de codage doit permettre la mise au jour des caractéristiques principales d'un objet étudié, et implique forcément une connaissance de celui-ci, donc des présupposés théoriques. Comme le fait très justement remarquer Encrevé (1992 : 110), « *le danger, c'est de faire croire que la représentation de l'oral, c'est de l'oral* ». Avoir conscience d'une telle singularité amène à se poser d'autres questions : quel est l'impact de ces présupposés dans une analyse ? Et surtout, peut-on se dégager de ces implicites pour traiter et analyser les données ?

Très certainement, le protocole du projet PFC ne peut éviter des positionnements différenciés entre description et théorie. Cependant, il apparaît nécessaire de les rendre explicites, de les évaluer et de savoir en mesurer les conséquences afin d'apporter une réponse possible aux objectifs de départ qui sont, rappelons-le, de (i) tester les théories existantes sur un objet, (ii) comparer les différentes variétés du français et enfin (iii) favoriser les échanges entre les connaissances.

5.2 Constitution et description du corpus

Nous avons décrit en 5.1., le cahier des charges et les techniques de constitution du corpus PFC. Dans cette section, nous dressons un bilan descriptif du corpus obtenu. Nous fournissons un certain nombre de données quantitatives qui permettent de se faire une représentation exacte à la fois de la diversité des données disponibles mais également de la portée du projet PFC.

5.2.1 Les points d'enquête

Le tableau ci-dessous (Figure 5-7) dresse un panorama de l'état d'avancement de tous les points d'enquêtes du projet PFC, présents et à venir, soit un total de 74 enquêtes (en septembre 2007). Chaque colonne indique l'état d'avancement des enquêtes :

Constitution et description du corpus

(i) terminées et disponibles sur le site (« en ligne »), (ii) en cours de transcription ou de codage (« en cours de traitement ») et (iii) à venir (« en projet »).

74 enquêtes						
		37 en ligne	27 en cours de traitement		9 en projet	
M I D I	1	Aix-Marseille (13)	38	Auriac-sur-Vendinelle (31)	66	Carcassonne (11)
	2	Béarn (64)	39	Cussac Fort-Médoc (33)		
	3	Biarritz (64)	40	Toulouse Centre Ville (31)		
	4	Douzens (11)				
	5	Lacaune (81)				
	6	Marseille Centre Ville (13)*				
	7	Nice (06)*				
	8	Rodez (12)				
	9	Salles-Curan (12)*				
	10	Toulouse (31)				
N O R D	11	Aveyronnais à Paris (75)	41	Amiens (80)	67	Briançon (05)
	12	Brecey (50)	42	Bar-sur-Aube (10)	68	Chambery (73)
	13	Brunoy (91)	43	Béthune (62)	69	Clermont-Ferrand (63)
	14	Cherbourg (50)*	44	Boersch (67)	70	Saint-Etienne (42)
	15	Dijon (21)	45	Domfrontais (61)	71	Strasbourg (67)
	16	Grenoble (38)	46	Haute Savoie (74)	72	Tours (37)
	17	Guadeloupéens à Paris (75)*	47	Houilles (78)		
	18	Ile de Sein (29)*	48	Joeuf (54)		
	19	Lyon (69)	49	Le Raincy (93)		
	20	Metz (57)*	50	Lille (62)		
	21	Nantes ville (44)	51	Mantes-la-Jolie (78)		
	22	Ogéville (54)	52	Marq-en-Bareuil (59)		
	23	Paris (Banlieue)*	53	Montargis (45)		
	24	Paris anc. Noblesse (75)	54	Mulhouse (68)		
	25	Puteaux (92)	55	Orléans (45)		
	26	Roanne (42)	56	Paris (75)		
	27	Treize-Vents (85)	57	Rouen (76)		
I N T E R N A T I O N A L	28	Abidjan (Côte d'Ivoire)	58	Acadie	73	Louisiane (Etats-Unis)
	29	Algérie *	59	Beyrouth (Liban)	74	Saint-Louis (Sénégal)
	30	Gembloux (Belgique)	60	Colombie Britannique		
	31	Ile de la Réunion	61	Genève alloglottes (Suisse)		
	32	Liège (Belgique)	62	Ile Maurice		
	33	Nyon (Suisse)	63	Martinique		
	34	Tournai (Belgique)	64	Nouvelle Calédonie		
	35	Alberta (Canada)*	65	Windsor (Ontario)		
	36	Québec city (université)*				
	37	Burkina Faso*				

Figure 5-7 Points d'enquêtes PFC

Les enquêtes sont réparties en différents groupes, en fonction de leur localisation géographique. La France (l'Hexagone) a été divisée en deux zones : le Nord et le Midi.

On obtient donc trois groupes : la zone Midi de l'Hexagone (**Midi**), la zone Nord de l'Hexagone (**Nord**) et la zone Internationale (**International**). Certains points d'enquêtes sont suivis d'un astérisque indiquant que l'enquête est terminée, mais que sa mise en ligne est en cours et sera effective au plus tard fin décembre 2008.

5.2.1.1 Couverture géographique

Le tableau ci-dessous (Figure 5-8) résume la répartition géographique des enquêtes et indique le nombre d'enquêtes achevées, en traitement ou en projet de chaque zone.

Trois zones ont donc été prises en considération : l'International, le Midi et le Nord. Au total, sur les 74 enquêtes, plus de la moitié (40 au total soit 54 %) correspondent à la zone Nord, près de 19 % (soit 14) à la zone Midi et 27 % (soit 20) à la zone International. La répartition sur l'ensemble du corpus est la suivante :

	Total	En ligne	En cours	En projet	%
International	20	10	8	2	27,02 %
Midi	14	10	3	1	18,91 %
Nord	40	17	17	6	54,05 %
Total	74	37	28	9	100 %

Figure 5-8 Répartition géographique des enquêtes

Pour résumer, 73 % des enquêtes couvrent la partie Hexagonale (Midi et Nord) et 27 % couvrent la zone International (Figure 5-9).

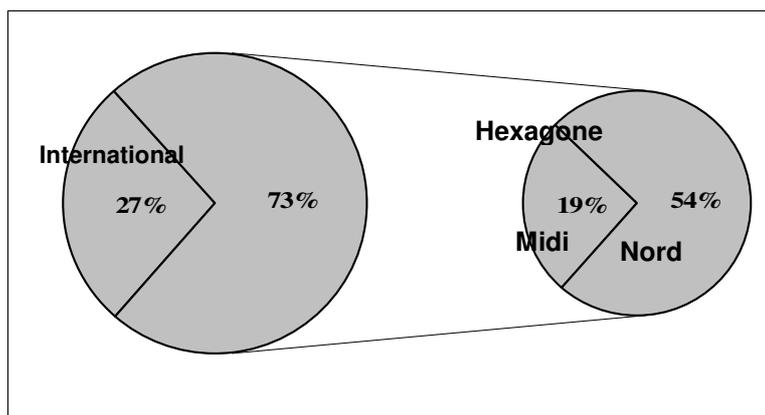


Figure 5-9 Représentation graphique de la répartition géographique des enquêtes

5.2.1.2 Etat d'avancement

A ce jour, sur les 74 points d'enquête, plus de 64 ont été couverts dans l'espace francophone international. 37 sont complètement terminés (codés, vérifiés, mis en ligne) et 27 sont en cours de traitement.

Le graphique suivant (Figure 5-10) illustre la répartition des enquêtes selon leur état d'avancement et montre que plus de la moitié des enquêtes prévues sont disponibles (51 %) et que près de 40 % sont en passe d'être achevées :

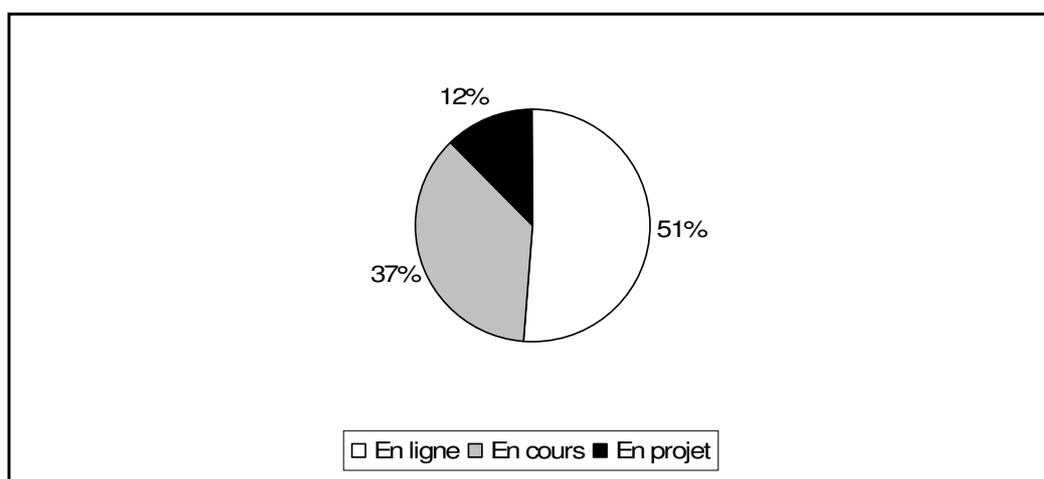


Figure 5-10 Représentation graphique de l'état d'avancement des enquêtes PFC

5.2.2 Les locuteurs

Concernant les informateurs PFC, les données que nous présentons ici ne correspondent pas à l'ensemble des enquêtes terminées que nous avons présentées en Figure 5-7.

Le nombre de locuteurs a été calculé en fonction des enquêtes disponibles sur le site en septembre 2007, soit un total de 24 enquêtes représentées dans le tableau ci-dessous (Figure 5-11). Les données qui sont analysées dans le Chapitre 6 (p. 175 et *sq.*) proviennent des enquêtes présentées dans cette section.

5.2.2.1 Répartition géographique

Le nombre de locuteurs par enquête s'échelonne entre 8 et 14. Au total, sur les 24 enquêtes, on compte 259 locuteurs interrogés, répartis géographiquement de la façon suivante (Figure 5-12) :

Zone	Enquêtes	Nombre de locuteurs
MIDI	Aix-Marseille	8
	Biarritz	12
	Douzens	10
	Lacaune	13
	Rodez	9
	Toulouse	14
Total		66
NORD	Aveyronnais à Paris	12
	Brécey	11
	Brunoy	10
	Dijon	8
	Grenoble	12
	Lyon	11
	Nantes	11
	Ogéville	12
	Paris Centre ville	12
	Puteaux-Courbevoie	6
	Roanne	9
	Vendée	8
	Total	
INTERNATIONAL	Abidjan	14
	Gembloux	12
	Ile de la Réunion	9
	Liège	12
	Nyon	12
	Tournai	12
Total		71

Figure 5-11 Répartition des locuteurs par enquête

La zone Nord comptabilise près de la moitié des locuteurs (au total 122, soit 45 %), l'autre moitié étant équitablement répartie entre le Midi (25 %) et la zone International (27 %). Cette répartition des locuteurs correspond à la répartition géographique des enquêtes observée en Figure 5-9, à savoir que l'Hexagone est beaucoup plus représenté que l'International : 73 % contre 27 %.

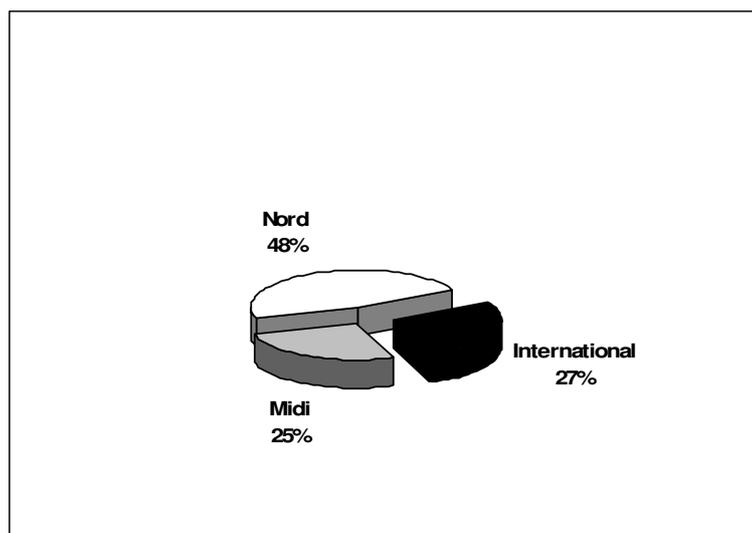


Figure 5-12 Représentation graphique de la répartition des enquêtés par secteur géographique

5.2.2.2 Sexe et âge des locuteurs

La répartition par sexe est à peu près égale. Les femmes représentent 53 % et les hommes 47 %.

Les locuteurs ont été répartis par le projet PFC en quatre tranches d'âge. La répartition se présente ainsi (voir aussi l'Annexe 6 pour le détail) :

Age	Interviewés
< 30 ans	94
30-45 ans	38
46-60 ans	57
> 60 ans	70
Total	259

Figure 5-13 Répartition des locuteurs par tranche d'âge

La Figure 5-14 montre que les témoins âgés de moins de 30 ans sont les plus nombreux. Ils représentent 37 % des enquêtés.

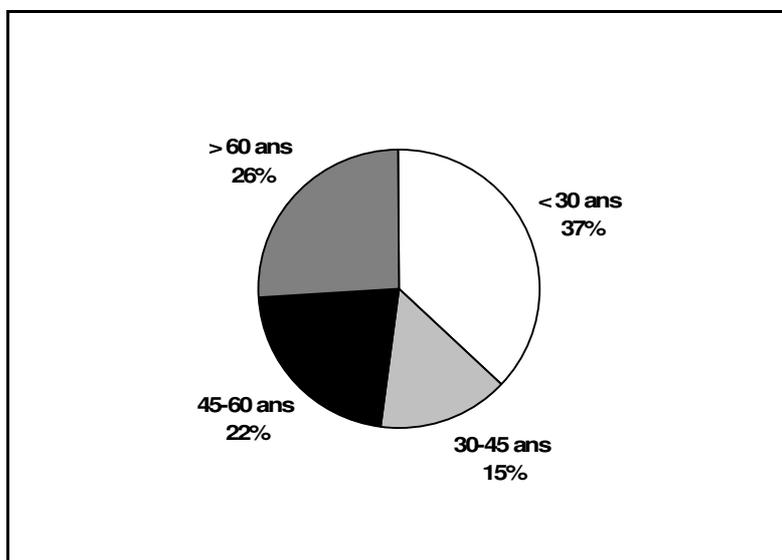


Figure 5-14 Représentation graphique de la répartition des témoins par tranche d'âge

Chapitre 6

Méthodologie, résultats et analyses

Dans ce chapitre, nous utiliserons les données extraites du corpus PFC pour tester un ensemble d'hypothèses que nous avons exposées dans le Chapitre 3.

Comme nous le verrons tout au long de ce chapitre, les résultats peuvent parfois être partiels, incomplets et ce, pour des raisons très diverses : soit méthodologiques, soit par manque d'informations nécessaires à l'hypothèse testée, soit encore parce que le nombre de contextes et / ou d'occurrences observés est trop important. Aussi, en introduction de chaque section, commencerons-nous par faire état de l'ensemble des problèmes apparus et de la méthodologie adoptée en conséquence. Nous présenterons ensuite l'ensemble des données retenues, puis les observations qui en découlent.

Dans un premier temps (*cf.* 6.1), nous commencerons par présenter les données de base qui ont été recueillies et par éliminer un certain nombre de celles qui, pour diverses raisons (nombre très faible d'occurrences, *etc.*), ne peuvent être retenues. Dans un second temps (*cf.* 6.2), nous débiterons les analyses en confrontant les données aux hypothèses formulées autour de la dimension sociolinguistique (le niveau stylistique, le niveau d'études et la classe d'âge). Dans un troisième temps (*cf.* 6.3), nous examinerons les différentes consonnes de liaison qui apparaissent dans le corpus ainsi que leur fréquence d'apparition, indépendamment de toute autre dimension. Cette analyse mettra en évidence un certain nombre de problèmes relatifs à la longueur. Aussi, dans un quatrième temps (*cf.* 6.4), ferons-nous le point sur la possible influence de la longueur dans la réalisation de la liaison. Enfin, dans un cinquième et dernier temps (*cf.* 6.5), nous examinerons une partie du contenu de la grille tripartite révisée de Delattre en fonction des contextes de notre propre corpus.

6.1 Introduction

Les résultats présentés dans ce chapitre s'appuient principalement sur les conversations guidée et libre fournies par PFC. Les résultats de la lecture de texte serviront plus particulièrement à une description sociolinguistique des réalisations de la liaison (*cf.* 6.2, pp. 188 et *sq.*) et à confirmer ou infirmer l'influence de certains paramètres, qui seront mis en évidence dans les analyses qui suivent, tels que la longueur du lieu, la construction syntaxique et la fréquence.

Toutes les données observées ici ont été extraites de la base de données PFC accessible en ligne (www.projet-pfc.net), grâce au moteur de recherche que le projet met à disposition. Elles proviennent de 24 enquêtes, soit un total de 259 locuteurs. Nous donnons dans le tableau suivant (*cf.* Figure 6-1), la liste de ces enquêtes (*Enquête*) par ordre alphabétique, ainsi que leur code enquête (*Code*) et le nombre de locuteurs que contient chacune d'elles (*Locuteurs*).

<i>Enquête</i>	Code	Locuteurs
Abidjan	cia	14 locuteurs
Aix-Marseille	13b	8 locuteurs
Aveyronnais à Paris	75x	12 locuteurs
Biarritz	64a	12 locuteurs
Brécey	50a	11 locuteurs
Brunoy	91x	10 locuteurs
Dijon	21a	8 locuteurs
Douzens	11a	10 locuteurs
Gembloux	bga	12 locuteurs
Grenoble	38a	12 locuteurs
Ile de la Réunion	974	9 locuteurs
Lacaune	81a	13 locuteurs
Liège	bla	12 locuteurs
Lyon	69a	11 locuteurs
Nantes	44a	11 locuteurs
Nyon	sva	12 locuteurs
Ogéviller	54b	12 locuteurs
Paris centre	75c	12 locuteurs
Puteaux-Courbevoie	92a	6 locuteurs
Roanne	42a	9 locuteurs
Rodez	12a	9 locuteurs
Toulouse	31a	14 locuteurs
Tournai	bta	12 locuteurs
Vendée	85a	8 locuteurs
24 enquêtes		259 locuteurs

Figure 6-1 Corpus d'enquêtes utilisées

L'ensemble de ces enquêtes rassemble 37 196 contextes¹⁰¹ codés pour la liaison. Ces contextes sont issus aussi bien de la lecture de texte que des deux types de conversations.

	Nombre de contextes codés
Liaisons non réalisées	18 797
Liaisons enchaînées	18 191
Liaisons incertaines	30
Liaisons épenthétiques	48
Liaisons non enchaînées	130
Nombre total de contextes	37 196

Figure 6-2 Nombre total de contextes « liaison » codés dans le corpus fourni par PFC

Les liaisons non réalisées et les liaisons enchaînées constituent les données qui sont au cœur de notre étude. Elles représentent 36 988 contextes de liaison. Il importe cependant d'examiner plus précisément ce qu'il en est des 3 autres cas (liaisons non enchaînées, liaisons incertaines et liaisons épenthétiques), pour lesquels PFC a prévu un codage particulier.

A l'issue de cette présentation, on aura justifié le fait que nous n'avons pas pris en compte ces contextes (*cf.* 6.1.1). On aura aussi fixé le nombre définitif de contextes qui constituent notre corpus d'étude (*cf.* 6.1.2) et que nous examinerons dans les sections suivantes :

- La liaison et les facteurs sociolinguistiques (*cf.* 6.2, pp.188 et *sq.*)
- Les consonnes de liaison (*cf.* 6.3, pp. 203 et *sq.*)
- L'influence de la longueur (*cf.* 6.4, pp. 260 et *sq.*)
- L'influence de la syntaxe (*cf.* 6.5, pp. 271 et *sq.*)

¹⁰¹ Au cours des analyses, nous utilisons indifféremment les termes de *contexte* et de *site*.

6.1.1 Les contextes de liaison non considérés

Nous reprenons dans le tableau ci-dessous l'ensemble des cas pour lesquels PFC a prévu un codage particulier :

	Nombre de contextes
Liaisons incertaines	30
Liaisons non enchaînées	130
Liaisons épenthétiques	48
Total	208

Figure 6-3 Liaisons non considérées

Dans les faits, ces trois cas ne se situent pas sur le même plan.

1. Le cas des liaisons incertaines

Les liaisons incertaines, au nombre de 30 (soit 0,08 % des contextes codés), ne peuvent en aucun cas être récupérées dans notre corpus d'étude : la mauvaise qualité sonore ne permet pas de déterminer précisément s'il y a eu liaison ou non.

2. Le cas des liaisons non enchaînées

Les liaisons non enchaînées sont au nombre de 130 (soit 0,35 % des contextes codés). Ces non-enchaînements sont le fait de 75 locuteurs (sur 256 au total) et se répartissent de la façon suivante :

- 58 non-enchaînements dans la conversation guidée ;
- 18 dans la conversation libre ;
- et 54 dans la lecture de texte.

Cependant, nous avons examiné l'ensemble de ces 130 occurrences : le non-enchaînement apparaît comme un phénomène très marginal dans les productions des locuteurs de PFC, d'autant que, dans les faits, un certain nombre de cas ne peuvent être considérés comme tels si l'on suit la définition d'Encrevé (*cf.* ci-dessous). En d'autres termes, PFC a une définition plus large du phénomène.

En effet, selon Encrevé (1988 : 31), les « *analyses empiriques* [sont] *indiscutables* ». Elles montrent qu'on peut « *entendre deux types clairement différenciés de réalisation phonétique correspondant respectivement à deux types de représentation phonétique* » :

- des liaisons enchaînées qui font apparaître la consonne de liaison en position d'attaque syllabique à l'initiale du mot qui suit la forme à laquelle cette consonne de liaison, dans sa forme graphique, est rattachée ;

- et des liaisons non enchaînées dans lesquelles la consonne de liaison est réalisée en position de coda syllabique à la finale de la forme à laquelle cette consonne de liaison est rattachée.

Encrevé (1988 : 35) ajoute que la principale caractéristique du non-enchaînement est qu'il y a « *quelque chose entre C et V qui concrétise la non-resyllabation de CL* ».

Notre position est la suivante : si le « *quelque chose* » d'Encrevé existe bien (coup de glotte ou réalisation de « h aspiré »), ce que nous ne remettons pas en cause, dans un certain nombre de non-enchaînements que nous avons relevés, ce qui se révèle particulièrement caractéristique est la présence de fortes hésitations. Ces hésitations sont attestées par les nombreuses reprises opérées par les locuteurs :

- 17 % pour la conversation guidée, soit 10 reprises ;
- 33 % pour la conversation libre, soit 6 reprises ;
- et 2 % pour la lecture du texte, soit une seule reprise¹⁰².

Nous utilisons intentionnellement le terme « d'hésitation » et pas celui de « pause-hésitation ». Il ne s'agit pas ici de « *pause-hésitation* » au sens de Coustenoble et Armstrong (1934), où le « locuteur cultivé » marque l'emphase grâce à l'insertion d'une « *plosive glottale* » en attaque de syllabe du mot lié, en lieu et place de la consonne à réaliser. Il ne s'agit pas non plus de « *l'accent d'insistance* » (Delattre, 1955 ; Ågren, 1973), ni de « *l'accent intellectif* » décrit par Fouché (1959), ni encore de pauses « *de repos* » ou « *d'arrêt* » (Lucci, 1983).

¹⁰² Si la lecture d'un texte à voix haute peut être considérée comme propice à la réalisation des liaisons non enchaînées, nous émettons cependant une réserve quant au faible pourcentage de « reprise correctrice » de ces liaisons non enchaînées. En effet, lors de la numérisation des enregistrements, les textes sont « nettoyés ». Les reprises correctives par exemple, sont supprimées des enregistrements, la première production étant seule retenue. Il est donc tout à fait possible, bien que nous n'en ayons aucune trace, que le nombre de reprises correctives contenues dans les lectures de texte soit plus élevé.

Il s'agit « d'hésitations » au sens propre. Le locuteur, toute catégorie socioprofessionnelle confondue, « cherche ses mots », hésite, produisant ainsi une liaison non enchaînée et finalement, se reprend pour produire la même liaison mais cette fois-ci enchaînée. Ces hésitations sont souvent marquées par la présence d'interjections (*ehh, enfin, etc.*), voire par l'intervention d'un interlocuteur qui interrompt le locuteur (exemple (34)).

Nous proposons ci-dessous quelques exemples de ces liaisons non enchaînées suivies d'une reprise corrective. Les liaisons non enchaînées sont signalées par (NE), les liaisons enchaînées par (LE) et les liaisons non réalisées par (NL). Les signes d'hésitations sont soulignés.

(31) *je pense qu'on (NE) enfin, je pense qu'on (LE) est beaucoup (31aab1, guidé)*

(32) *un grossiste s'est (NE) ehh, s'est (LE) installé sur Roanne (42aml1, guidé)*

(33) *y a quand même un, un (NE), un (LE) aveyronnais (75xrg1, guidé)*

(34) *ça dépend ce que vous (NE) <E : (XX)>, ça dépend ce que vous (LE) appelez (81aar1, guidé)*

(35) *c'est (NE), c'est (LE) un autre, c'était (NL) une autre école (12atp1, libre)*

(36) *ehh, on s'en (NE), on s'en (LE) est servi (91adb1, libre)*

(37) *comme tous les (NE), tous les (LE) autres enfants hein (91ael1, libre)*

(38) *Des (LE) activiti/, des (NE), des (LE) activistes (92acd1, texte)*

D'après nos estimations, 20 à 30 % des liaisons identifiées au départ comme « non enchaînées » dans les conversations libres et guidées ne constituent pas de « vrais » cas de non-enchaînement au sens d'Encrevé.

Cette très faible représentation du phénomène de non-enchaînement tend à prouver qu'il est principalement réservé aux spécialistes de la parole publique

(cf. Encrevé, 1988). Ces propos nécessiteraient toutefois une vérification auprès de locuteurs exerçant des professions pour lesquelles le discours public est indissociable de leur fonction.

3. Le cas des liaisons épenthétiques

Les liaisons épenthétiques sont au nombre de 48 (soit 0,13 % des contextes codés). Cependant, à l'intérieur de cette catégorie des « épenthétiques », il est nécessaire de distinguer 3 cas :

- a) 14 épenthèses (au sens de PFC), soit 0,037 % des contextes codés ;
- b) 6 liaisons avec CL différente (soit 0,016 % des contextes codés), c'est-à-dire des liaisons enchaînées dans lesquelles la consonne réalisée n'est pas le pendant phonétique de la consonne présente dans la forme graphique, mais qui cependant ne correspondent pas aux critères posés par PFC pour la liaison réalisée avec épenthèse ;
- c) et 28 erreurs de codage, soit 0,075 % des contextes codés.

a. Les épenthèses

Le projet PFC code comme épenthétiques les liaisons dont la forme liante ne présente aucune consonne graphique de liaison en finale. Tous registres confondus (la lecture de texte et les deux types de conversations), on ne dénombre qu'un taux infime de ces liaisons sur l'ensemble du corpus : au total, quatorze cas. Les consonnes épenthésées sont au nombre de deux : [z] et [t], c'est-à-dire deux des consonnes de liaison considérées comme réalisant le plus de liaisons (cf. 3.1.2.1).

Les cas de liaisons épenthétiques se concentrent avant tout dans la lecture du texte : treize cas au total, dont deux correspondent à l'insertion d'un [z] après *quatrième* (exemple (39)), dix à l'insertion d'un [t] après la forme verbale *mène* (exemple (40)) et un à l'insertion de [t] après la forme verbale *a* (exemple (41)).

(39) *quatrième [z] aux Jeux Olympiques (12ajp1 et 69acg1)*

(40) *qui mène [t] au village (13amb1, 13bma1, 31alm1, 31aoe1, 50apb1, 75ccr2, 75csb1, 92acd1, 92acj1 et abafb1)*

(41) *le Premier Ministre a [t] en effet décidé (50apb1)*

La conversation guidée présente quant à elle un seul cas d'épenthèse (exemple (42)), dans lequel le locuteur insère un [z] après la forme pronominale singulier *il* qui contient une consonne finale fixe [l].

(42) *il [z] a (13aog1)*

b. Les liaisons avec CL différente

Au total, nous avons relevé 6 contextes de liaison dans lesquels le locuteur réalise une consonne de liaison différente de la consonne attendue (*i.e.* différente de la consonne graphique présente dans la forme lexicale du lieu). Ces productions particulières sont le fait de 6 locuteurs différents et constituent un cas particulier. En effet, le codage PFC ne prévoit pas de distinction entre liaison enchaînée avec « CL phonétique corrélée à CL graphique » et liaison enchaînée avec « CL phonétique non corrélée à CL graphique » (*cf.* 5.1.3, pp. 159 et *sq.*). De ce fait, les codeurs ont noté ces liaisons comme des liaisons épenthétiques¹⁰³.

Quatre cas peuvent être observés dans la conversation libre (exemples (43), (44), (45) et (46)). Sur trois de ces contextes de liaison, on observe une tentative de « réparation » de la part du locuteur (exemples (43), (45) et (46)). Se rendant compte de son « erreur », ce dernier reprend son énoncé pour le corriger et réaliser la consonne de liaison attendue, autrement dit la consonne de liaison présente morphologiquement dans le mot :

(43) *des anciens [n] employés, des anciens [z] employés (21acp1, libre)*

(44) *mettre cent [z] euros (blajv1, libre)*

(45) *les immeubles de bas [t] en haut, de bas [z] en haut (69amg1, libre)*

(46) *puis finalement, en [t], en [n] allant (91aell, libre)*

Deux autres cas sont observés dans la lecture du texte (exemples (47) et (48)) :

¹⁰³ Ces 6 cas de liaisons apparaissent clairement lorsque l'on se penche sur les liaisons codées comme épenthétiques et qui sont en nombre restreint. Il est cependant possible que d'autres liaisons de ce type aient été codées comme des liaisons « simplement » enchaînées et qu'elles soient intégrées à notre corpus (*cf.* 6.1.2) sans que nous puissions les dénombrer avec exactitude.

(47) *provoquer [t] une explosion (75xlv1, texte lu)*

(48) *un grand [n] émoi (50aid1, texte lu)*

Parmi ces exemples, on remarque une erreur de type *morphologique* (cf. 3.3.4), l'exemple (44), dans lequel le locuteur a réalisé une consonne de liaison ([z]) indicatrice du pluriel.

L'exemple (45) illustre probablement une erreur de type *lexical*, dans laquelle le [t] de *bas [t] en haut* est réalisé par analogie avec celui réalisé phonétiquement dans l'expression *de haut [t] en bas*.

Les exemples (46) et (48) illustrent une erreur à *distance*. En (46), le locuteur réalise une liaison avec la consonne de liaison de *finalement [t]* au lieu d'utiliser celle de *en [n]*, tandis qu'en (48), la liaison est réalisée avec le déterminant *un [n]* qui précède l'adjectif *grand [t]*, lui-même suivi du nom *émoi*.

c. Les erreurs de codage

Enfin, nous avons relevé 28 erreurs de codage parmi les contextes codés comme étant des liaisons épenthétiques. Ces erreurs sont de trois types :

(i) Au lieu de coder la consonne de liaison réalisée devant consonne comme le recommande le projet PFC (*i.e.* 1KC, cf. 5.1.3.2), le codeur note une liaison épenthétique. Dans l'exemple (49), le codage aurait dû être de la forme *ItC* soit *quand1ItC* :

(49) *quand14t c'est des familles de jeunes (50ajm1)*

(ii) Un autre type d'erreur consiste à coder comme épenthétique une liaison réalisée avec enchaînement alors qu'il n'y a pas d'épenthèse. Dans les exemples (50) et (51), le codage aurait dû prendre la forme *12t*, au lieu de *14t* :

(50) *est en grand14t émoi (21acl1, 42amg1, 50aid1)*

(51) *ce grand14t honneur (21acl1, 42amg1)*

(iii) Le dernier type d'erreur a une double origine. Il est la somme d'une erreur de type (ii) et d'une transcription orthographique erronée.

(52) *a quatre-vingt14z employés (50alb1)*

Dans cet exemple, la consonne de liaison [z] a été codée comme épenthétique, alors qu'une autre consonne de liaison ([t]) est présente graphiquement en finale du lieu (*quatre-vingt*). D'autre part, il s'avère que cette consonne de liaison présente morphologiquement dans la forme de *quatre-vingt* n'est pas celle que recommande la norme dans un emploi déterminatif, à savoir un *s*¹⁰⁴. Le codage de cette occurrence devrait signaler qu'il y a réalisation d'une liaison en [z] avec enchaînement, soit *quatre-vingts11z employés*.

4. Bilan

Nous venons de présenter les trois cas de liaison qui ne seront pas exploités dans les analyses que nous proposerons dans les sections suivantes : les liaisons incertaines, les liaisons non enchaînées et les liaisons épenthétiques qui nécessitent elles-mêmes une sous-catégorisation. Nous considérons en outre ces liaisons comme étant marginales du fait même de leur très faible représentation sur l'ensemble des contextes codés (0,56 %).

6.1.2 Les contextes de liaison considérés

Après avoir exclu les contextes de liaisons incertaines, non enchaînées et épenthétiques, le nombre total de contextes considérés dans notre corpus d'étude s'élève à 36 988.

¹⁰⁴ Les mots grammaticaux tels que les adjectifs numéraux cardinaux ont un comportement semblable aux autres déterminants. Ils sont construits à partir d'une liste fermée (*un, deux, vingt, cent, mille, etc.*). Les formes simples peuvent se combiner pour donner des formes construites comme par exemple lorsqu'il y a multiplication des éléments (par ex. *quatre-vingts, deux cents, etc.*). En dehors de cet emploi déterminatif, les adjectifs numéraux cardinaux peuvent se comporter comme des noms (par ex. *Tous les cinq sont là*) mais également comme des ordinaux. Dans ce dernier cas, ils ne s'accorderont ni en genre, ni en nombre (par ex. *la page quatre-vingt*). Voir également Grevisse (1969 : 894, art. 577b.).

	Nombre de contextes codés
Liaisons non réalisées	18 797
Liaisons enchaînées	18 191
Nombre total de contextes	36 988

Figure 6-4 Contextes de liaison considérés

Ces 36 988 contextes de liaison sont extraits des différentes situations proposées par PFC :

- 8 469 contextes proviennent de la lecture du texte ;
- 28 519 contextes proviennent des deux types de conversations et se répartissent ainsi : 15 345 contextes dans la conversation guidée et 13 174 contextes dans la conversation libre.

Nous verrons par la suite que la tripartition nécessite une sérieuse remise en question. Cependant, à cette étape de l'exposé, nous conservons cette terminologie pour rendre compte des résultats.

Par ailleurs, sur le plan de la constitution du corpus, il est nécessaire de garder en tête que nous ne disposons pas de l'ensemble des informations puisque, contrairement aux liaisons « variables » et « catégoriques » dont on connaît tous les contextes, qu'ils soient réalisées ou non, on ne connaît des liaisons « erratiques » que celles qui sont réalisées.

Le tableau ci-dessous (*cf.* Figure 6-5) donne le nombre total de contextes codés (*Nb de contextes*) pour la conversation (guidée et libre) et la lecture du texte. Il fournit également les pourcentages de réalisation de la liaison de chacun des deux volets (*% de L*).

On observe, sur la base de la fréquence globale des liaisons, une nette opposition entre les deux types de situation : le pourcentage de liaisons effectives est plus faible dans la conversation (45 %) que dans la lecture du texte (63,2 %). Il semble donc qu'une situation de lecture ait une influence positive sur l'accroissement des fréquences de réalisation de la liaison.

	Nombre de contextes	% de L
Conversation	28 519	45 %
Lecture de texte	8 469	63,2 %

Figure 6-5 Réalisation de la liaison en lecture et en conversation

Ces résultats doivent toutefois être nuancés. En effet, on remarque que le nombre de contextes présents dans la conversation est près de quatre fois plus important que celui de la lecture de texte, ce qui peut avoir une influence sur les pourcentages de réalisation de la liaison. Si les contextes les plus représentés dans la conversation correspondent à des liaisons variables très peu fréquemment réalisées (par ex. Nom pluriel +) et non représentées dans le texte, les résultats risquent d'être faussés. Cependant, à défaut d'estimer¹⁰⁵, on peut espérer que tel n'est pas le cas, puisque le texte a été conçu de telle façon qu'il recouvre un large échantillon des différents contextes possibles (*cf.* Annexe 8).

Un dernier paramètre est à considérer : les contextes de liaison présents dans la lecture de texte sont similaires¹⁰⁶ à tous les locuteurs, ce qui n'est pas le cas des contextes de la conversation, puisque dans cette situation de communication, on ne peut contrôler ni prédire aucun des contextes possibles. Pour ces différentes raisons, il importe de rester prudent quant aux conclusions et hypothèses que l'on pourrait formuler.

Dans un premier temps (*cf.* 6.2), nous examinons les réalisations de la liaison en fonction de facteurs sociolinguistiques.

¹⁰⁵ Il nous est en effet impossible, à l'heure actuelle, de faire un compte rendu exact des quelques 28 519 contextes de liaison codés dans la conversation que nous pourrions comparer avec les contextes présents dans le texte (8 469 au total).

¹⁰⁶ Similaires ou presque si l'on tient compte du fait que (i) nous avons éliminé un certain nombre de sites potentiels et que (ii) il est toujours possible que les codeurs aient, ou bien, noté des contextes qui n'auraient pas dû l'être (par. ex. des liaisons décrites comme erratiques et qui ne sont pas réalisées), ou bien, oublié un certain nombre de sites prévus au codage (par ex. des liaisons réalisées). Le texte comprend minimalement 32 contextes à coder et maximalement 35 (avec les liaisons erratiques réalisées, au nombre de 3). Sur 259 locuteurs, on peut donc espérer un minimum de 8 288 contextes codés et un maximum de 9 065.

6.2 La liaison et les facteurs sociolinguistiques

Dans cette section, nous analysons la fréquence de la liaison en fonction des facteurs sociolinguistiques.

Nous adoptons ici la définition de « style » au sens de Gadet (2003 : 2)¹⁰⁷ :

Par « style », on entend la propriété des langues selon laquelle, partout pour autant que l'on sache, les usagers des langues disposent de répertoires diversifiés dans leur langue. Cette diversification intervient soit en fonction des circonstances, de la situation et des participants à l'échange (style situationnel, en réponse à certaines composantes d'une situation), soit selon les ressources déployées dans les interactions (style initiatif, permettant au locuteur de redéfinir partiellement une situation).

Etant donné le protocole mis en place par le projet PFC (*cf.* 5.1.1), nous utilisons le terme de « style » dans le sens de « style situationnel ».

En premier lieu, nous observons les taux de réalisation de la liaison en fonction du style utilisé, indépendamment de tout autre critère (*cf.* 6.2.1). Ensuite, nous incluons deux autres paramètres sociolinguistiques : le niveau d'études des locuteurs (*cf.* 6.2.2) et la tranche d'âge à laquelle ils appartiennent (*cf.* 6.2.3).

6.2.1 Fréquence de la liaison en fonction du style

Les spécialistes de la liaison sont unanimes sur le fait que la fréquence de la liaison varie selon les différents niveaux d'usage de la langue (Delattre, 1955 et 1966 ; Ågren, 1973 ; Encrevé, 1988 pour ne citer qu'eux) : la situation dans laquelle se déroule un entretien joue un rôle dans la nature des composantes linguistiques. Un même locuteur utilise différents registres de langue selon la situation¹⁰⁸. Plus le style utilisé est soutenu, plus les pourcentages de réalisation de la liaison ont tendance à croître (*cf.* 3.3.3.1, pp. 87 et *sq.*).

¹⁰⁷ *Cf.* également 3.3.1.2, pp. 83 et *sq.*

¹⁰⁸ Labov (1972a) remarque toutefois que l'écart entre la langue courante et la langue soutenue est beaucoup plus marqué chez les populations les moins scolarisées.

Afin de pouvoir comparer les taux de réalisation de la liaison en fonction du style, nous avons relevé (*cf.* Figure 6-6), pour chaque type de conversation ainsi que pour la lecture de texte, le nombre total de contextes de liaison pour lesquels nous avons identifié le nombre total de liaisons réalisées et de liaisons non réalisées. Nous avons également calculé les pourcentages correspondants aux liaisons réalisées (*% de L*) et aux liaisons non réalisées (*% de NL*) sur l'ensemble des contextes codés.

	Nombre total de contextes	Liaisons réalisées		Liaisons non réalisées	
		Nombre total	% de L	Nombre total	% de NL
Guidé	15 345	7 037	45,9 %	8 308	54,1 %
Libre	13 174	5 804	44,1 %	7 370	55,9 %
Texte	8 469	5 350	63,2 %	3 119	36,8 %

Figure 6-6 Réalisation de la liaison selon le registre

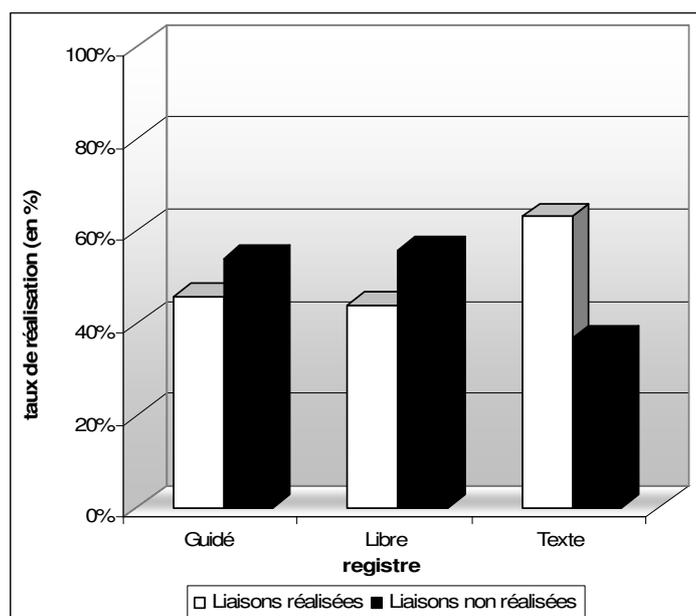


Figure 6-7 Représentation graphique de la réalisation de la liaison selon le registre

1. Différence entre conversations et lecture

La réalisation de la liaison se caractérise globalement par une opposition entre la situation de lecture et la situation conversationnelle (*cf.* Figure 6-7 pour une visualisation par histogramme) : les liaisons sont plus réalisées lorsque le témoin se

trouve en situation de lecture (63 %) qu'en situation conversationnelle (45 % pour les deux types de conversation confondus).

2. Pas de différence entre les deux types de conversation

On peut aussi constater que les taux de réalisation de la liaison diffèrent légèrement d'un type de conversation à un autre. Mais la différence (45,9 % de liaisons réalisées dans la conversation guidée et 44,1 % dans la conversation libre) est infime. Ce résultat peut surprendre lorsque l'on considère l'hypothèse, très largement partagée et reprise dans la littérature, qu'une situation formelle favorise la liaison chez les locuteurs (*cf.* 3.3.3.1). Rappelons que Delattre (1955) est un des premiers auteurs à faire état de l'influence du facteur stylistique sur la fréquence de réalisation de la liaison (*cf.* citation p. 98) :

Le facteur stylistique est de beaucoup le plus fort. Le choix du style, exigé par les circonstances, peut faire varier considérablement la fréquence des liaisons. « A mesure que le ton s'abaisse et devient plus familier, » dit Pierre Fouché [1935]. (Delattre, 1955 : 44)

Delattre (1955 : 45) ajoute cependant que la formalité de la situation va de pair avec le profil sociolinguistique des locuteurs : « *ce qui précède [cf. citation] s'applique au langage de la classe des plus cultivés. A mesure que l'on s'éloigne de cette classe, le nombre de liaisons diminue* ».

Dans l'état actuel des résultats présentés, nous ne pouvons conclure. Il nous faut d'abord examiner les résultats en fonction du niveau d'études des locuteurs et de leur âge (*cf.* sections suivantes). Nous avons bien conscience de réduire la portée de Delattre en identifiant niveau d'études et « classe des plus cultivés », mais le niveau d'études est la seule donnée dont nous disposons qui permet d'objectiver, certes très partiellement, la notion de « classe cultivée ».

6.2.2 Fréquence de la liaison en fonction du niveau d'études

Le tableau ci-dessous (*cf.* Figure 6-8) précise les pourcentages de liaisons effectives (% L) dans les trois registres (*Libre, Guidé, Texte*) selon le niveau d'études atteint par

les locuteurs (*Niveau d'études*). Cependant, tous les locuteurs n'ont pas spécifié aux enquêteurs leur niveau d'études. En conséquence, nous ne pouvons présenter ici qu'un sous-ensemble de ces locuteurs. Au total, sur les 259 locuteurs répertoriés dans la base, seuls 147 d'entre eux sont référencés.

Ces locuteurs se répartissent de la façon suivante :

- 54 locuteurs ont arrêté leurs études en primaire ou dans les premières années de collège ;
- 28 locuteurs ont arrêté leurs études avant le lycée ;
- 13 locuteurs ont arrêté leurs études à la fin du lycée ;
- 7 locuteurs ont arrêté leurs études entre le Bac et le Bac+2 ;
- et 45 locuteurs ont atteint un niveau d'études supérieur à Bac+2.

Niveau d'études	Libre % de L	Guidé % de L	Texte % de L
Primaire / début collège	43,5 %	47,3 %	65,2 %
Collège	53,6 %	53,8 %	64,4 %
Lycée	43,7 %	44,4 %	63,3 %
Bac à Bac + 2	37,6 %	40,5 %	56,5 %
Bac + 2 à Bac + 4 ans	46,7 %	48,5 %	65,8 %
Bac + 4 et plus	42,1 %	41,7 %	61,6 %

Figure 6-8 Réalisation de la liaison selon le niveau d'études

Ces données sont représentées sous forme graphique dans la Figure 6-9 :

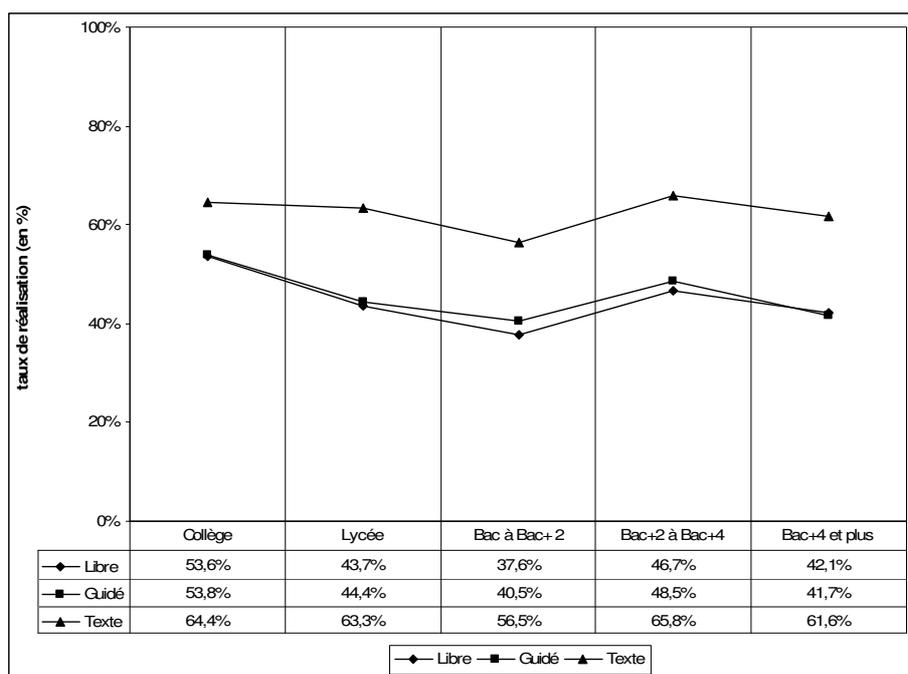


Figure 6-9 Représentation graphique de la réalisation de la liaison selon le niveau d'études

Le fait de ne pas avoir disposé de l'ensemble des données concernant les locuteurs fragilise les différentes hypothèses que l'on pourrait inférer de l'analyse des résultats, notamment celles qui chercheraient à établir les causes de la différence de comportement de telle ou telle classe. Par contre, il nous semble possible de discuter des résultats plus globalement.

1. Pas de relation entre le nombre de liaisons réalisées et le niveau d'études

L'observation de la courbe présentée dans la figure 6-9 indique que celle-ci ne croît pas avec le niveau d'études. Les collégiens réalisent même plus de liaisons que toutes les autres classes. Il y a bien des différences, mais elles restent extrêmement faibles.

2. Différence entre conversations et lecture

On retrouve bien la différence entre lecture et conversations, sans pour autant pouvoir distinguer la conversation libre de la conversation guidée.

3. La situation de lecture est celle qui produit le plus de liaisons

La relation lecture / liaison demeure robuste quel que soit le niveau d'études. Dans l'ensemble, le taux de réalisation de la liaison dans la situation de lecture est relativement régulier et élevé. Il est toujours au-dessus de 55 %.

L'absence de relation observée entre le nombre de liaisons et le niveau d'études des locuteurs peut paraître surprenante. Il entre très directement en contradiction avec l'hypothèse selon laquelle ce sont les locuteurs les plus scolarisés qui réalisent le plus de liaisons (*cf.* Gadet, 1989 : 52-53).

Cependant, l'éventuelle remise en question de cette hypothèse nécessite des études approfondies par le biais d'outils PFC qui permettraient (i) d'identifier à l'intérieur même des liaisons réalisées, les liaisons catégoriques des liaisons variables, et (ii) de confronter les taux de liaisons non réalisées, qui sont par définition variables, aux taux de liaisons variables réalisées et ce, pour chaque niveau d'études.

6.2.3 Fréquence des liaisons en fonction de la classe d'âge

Si l'on examine maintenant la fréquence de la liaison en fonction des tranches d'âge des locuteurs, on s'aperçoit que la réalisation de la liaison semble être proportionnelle à la tranche d'âge.

Le tableau ci-dessous (Figure 6-10) présente les taux de réalisation de la liaison calculés pour chaque tranche d'âge et selon le registre employé. Nous avons identifié les 4 tranches d'âge suivantes :

- les locuteurs âgés de moins de 30 ans au moment de l'enquête (<30 ans), au nombre de 94 ;
- ceux compris entre 30 et 45 ans (30-45 ans), au nombre de 38 ;
- les 46-60 ans (46-60 ans), au nombre de 57,
- et les locuteurs âgés de plus de 60 ans (>60 ans), au nombre de 70.

Tranche d'âge	Libre	Guidé	Texte
	% de L	% de L	% de L
< 30 ans	41,7 %	43,4 %	62,2 %
30-45 ans	40,6 %	44,9 %	61,2 %
46-60 ans	45,7 %	46,6 %	64,3 %
> 60 ans	50 %	51,7 %	65,9 %

Figure 6-10 Réalisation de la liaison selon la classe d'âge

Ces mêmes données sont synthétisées sous une autre forme dans la Figure 6-11 :

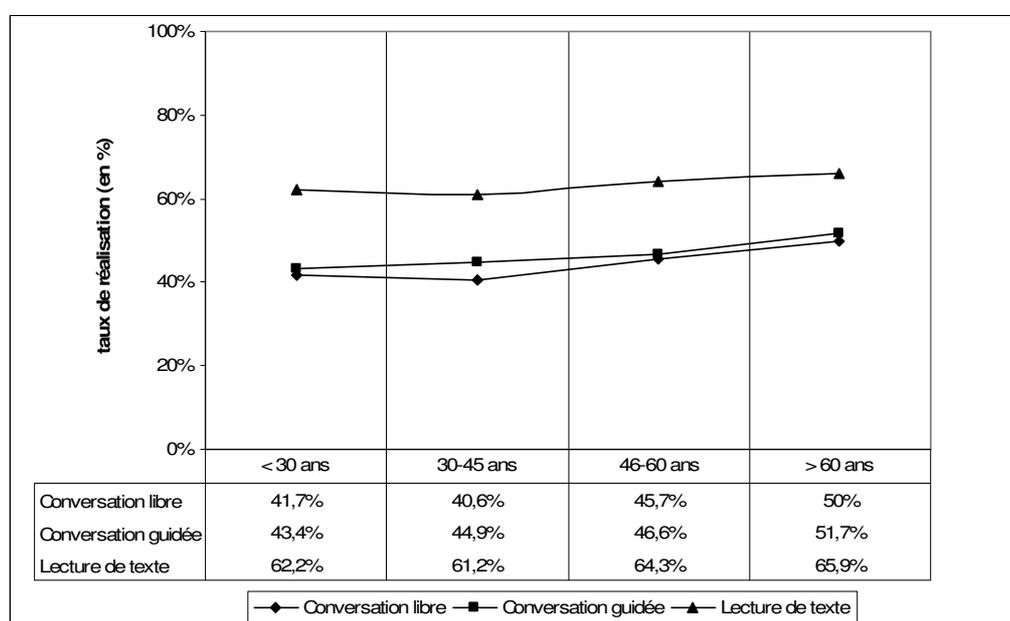


Figure 6-11 Représentation graphique de la réalisation de la liaison selon la classe d'âge

Globalement, les taux augmentent au fur et à mesure que l'âge des locuteurs augmente. Les plus de 60 ans présentent les taux de réalisation les plus forts : 50 % pour la conversation libre, 51,7 % pour la conversation guidée et 65,9 % pour la lecture de texte (cf. Figure 6-11).

1. Les locuteurs les plus âgés réalisent le plus de liaisons en situation conversationnelle

La réalisation de la liaison est plus importante à mesure que l'âge des locuteurs augmente. Là où les plus de 60 ans affichent des taux de réalisation supérieurs à 50 %,

les générations plus jeunes ne voient pas leurs taux dépasser les 47 % : entre 41,7 % et 43,4 % pour les moins de 30 ans ; entre 40,6 % et 44,9 % pour les 30-45 ans et entre 45,7 % et 46,6 % pour les 46-60 ans.

2. La situation de lecture estompe les différences générationnelles

C'est dans la situation de lecture que les taux de réalisation de la liaison les plus importants sont observés (*cf.* section précédente 6.2.1). Cependant, l'écart observé entre les différentes générations est beaucoup plus restreint que dans les situations de conversation. Les taux sont tous supérieurs ou égaux à 61,2 % et ne dépassent pas les 65,9 %.

6.2.4 Bilan

Dans cette section, nous avons tout d'abord vu et par la suite vérifié que les résultats PFC ne présentent qu'une différence infime sur le plan de la fréquence de la liaison entre la conversation guidée et la conversation libre et que la seule opposition pertinente porte sur l'examen comparé de la conversation dans sa globalité et de la lecture du texte. Ensuite, nous avons établi que, de façon plus générale (tous registres confondus), la fréquence de la liaison semble être corrélée avec la tranche d'âge des locuteurs mais pas avec leur niveau d'études.

La première question à laquelle nous essayons de répondre ici est de comprendre pourquoi la différence entre les deux registres de conversation est aussi faible, d'étudier plus précisément ce qui se passe sur le plan diaphasique. La seconde question porte sur le fait que l'importance de la dimension diastratique ne soit pas confirmée.

6.2.4.1 Variation diaphasique

Tous les auteurs s'accordent sur le principe que les liaisons sont plus fréquemment réalisées dans une situation formelle (*cf.* 3.3.3 pp. 87 et *sq.*). Par liaison, nous entendons ici « liaisons variables ». Autrement dit, les liaisons variables sont plus souvent réalisées en situation formelle qu'en situation informelle, puisque « *la liaison [variable] dépend*

du style. Elle se fait d'autant moins que le style est plus familier » (Delattre, 1947 : 149) (cf. 3.3.3.1, pp. 88 et sq.).

Comment alors interpréter cette quasi absence de différence de réalisation de la liaison entre les conversations guidée et libre ?

1. Première hypothèse : incapacité de PFC à générer deux registres distincts

La première hypothèse envisagée est celle de l'incapacité de PFC à générer deux registres de parole distincts : un registre soutenu et un registre familier.

Dans la plupart des cas, lors des interviews PFC, la conversation guidée et la conversation libre sont consécutives. On pourrait alors très bien concevoir que ces deux conversations n'en constituent en réalité qu'une seule. Les résultats obtenus pourraient alors s'interpréter de deux façons :

- soit ces taux qui, en moyenne, s'échelonnent entre 41 % et 45 %, représentent ce que l'on peut espérer obtenir de mieux en termes de « fréquence d'apparition » de la liaison dans une conversation informelle (conversation libre) ;

- soit ils représentent les pourcentages maximum de la réalisation de la liaison et correspondent en fait plus à des taux de réalisation proches du style formel (conversation guidée).

Les taux constitueraient alors approximativement les pourcentages les plus élevés que l'on puisse obtenir ou bien en situation de conversation informelle ou bien en conversation formelle.

L'une des volontés de PFC est d'éviter au maximum le paradoxe de l'observateur. On pourrait concevoir que cet objectif ait parfaitement fonctionné. Le locuteur, conscient au début de l'entretien d'être dans une situation de « surveillance », relâche peu à peu sa vigilance au fur et à mesure que la conversation avance, qu'il narre ses expériences et échange des idées avec son ou ses interlocuteurs. Il opérerait alors « inconsciemment » pour un style informel. Mais comme nous l'avons déjà dit, l'inverse nous paraît également possible, à savoir que le locuteur a pu rester sur son impression de « test » et continuer à surveiller ses productions orales tout au long de l'entretien. Le locuteur opérerait alors « consciemment ou non » pour un style formel.

Cette hypothèse est toutefois mise à mal par le fait que les données obtenues proviennent de 24 enquêtes différentes et menées par tout autant d'enquêteurs.

Tenir jusqu'au bout la position selon laquelle plus d'une vingtaine d'enquêteurs, opérant auprès de 259 locuteurs durant de plus de 300 heures d'interviews, n'aient pas pu obtenir les deux styles de conversation, ne nous semble pas non plus tenable. Cela reviendrait à affirmer qu'une grande majorité de nos locuteurs n'ont pu oublier leur situation d'enquêtés ou inversement qu'une majorité d'entre eux a trop vite oublié la présence de l'enregistreur et sa situation d'enquêté.

Nous avons vu (*cf.* pp. 145 et *sq.*) que le protocole PFC recommande que :

- (i) les entretiens guidés soient menés par l'enquêteur le moins proche du témoin sous forme de questions ;
- (ii) et que les entretiens libres soient une conversation de groupe permettant à l'enquêté d'oublier la présence de l'appareil enregistreur et ce, afin de toucher au plus proche de son vernaculaire.

Ayant nous-même participé à un certain nombre d'interviews, nous sommes consciente de la difficulté à respecter scrupuleusement ces exigences, mais l'expérience nous a aussi montré que l'enquêté est bien souvent beaucoup plus « à l'aise » lorsque les conversations sont menées en groupe (conversation libre). Il n'est généralement pas le seul intervenant, les discussions sont très variées, elles ne se concentrent pas seulement sur « l'identité du locuteur »¹⁰⁹ et ne portent sur aucun sujet prédéfini (sport, vacances, école, *etc.*).

Aussi, nous pensons qu'il est effectivement nécessaire de reconnaître :

1. que la plupart des enquêteurs ont respecté les exigences du protocole PFC ;
2. et que le nombre des interviews ne répondant pas aux critères nécessaires à la mise en place d'une distinction forte entre la conversation guidée et la conversation libre doit être insuffisant pour influencer sur la totalité des résultats présentés, manipulés selon des paramètres distincts.

¹⁰⁹ Contrairement à la conversation guidée qui permet notamment d'enrichir la fiche signalétique du locuteur (*cf.* Annexe 3).

2. Seconde hypothèse : les deux registres obtenus ne correspondent pas aux registres souhaités et la liaison ne constitue pas un indicateur permettant de les distinguer

Si l'on considère comme exacte (i) la proposition selon laquelle le taux de réalisation de la liaison variable est plus élevé dans un contexte formel que dans un contexte informel et que l'on tient pour acquis (ii) que la mise en application de la méthodologie adoptée par PFC a permis de produire deux types de conversations qui correspondent à des registres différents, alors on peut se demander si les registres générés sont bien ceux que l'on cherchait à obtenir.

Nous avons vu que :

- la lecture de texte, la conversation libre et la conversation guidée visent à obtenir trois registres différents (*cf.* Figure 5-1 : 148) : le registre conférence, la conversation soignée (ou style soutenu selon la terminologie employée) et le registre familial ;

- il y a consensus sur la différence de réalisation de la liaison entre un registre formel et un registre informel ;

- enfin, le taux de réalisation de la liaison est effectivement plus important dans la lecture du texte que dans les deux types de conversations.

Aussi, l'on pourrait considérer que la conversation libre correspond bien au registre « familial » identifié à la fois par la grammaire traditionnelle et par Delattre (1966), mais que la conversation guidée ne correspond pas au registre « soigné » / « soutenu » mais plutôt au registre « standard » que PFC ne cherchait pas à obtenir.

Le tableau Figure 5-1 prendrait alors la forme suivante :

	Classification traditionnelle des niveaux de langue	Fouché (1959)	Delattre (1966)	PFC répartition des corpus
ORAL / ECRIT	/	/	récitation de vers	lecture de texte
	/	/	conférence	
ORAL	soutenu	soutenu	soigné	/
	standard	courant	familier	conversation guidée
	familier			conversation libre
	populaire	/	/	/

Figure 6-12 Répartition des corpus PFC selon les registres

Si cette répartition des corpus PFC est correcte, la distinction, non significative, entre les conversations libre et guidée, pourrait s'expliquer par le fait que la liaison ne constitue pas un indicateur permettant de distinguer le registre standard du registre familier.

3. Troisième hypothèse : les deux registres obtenus sont bien les registres visés mais la liaison ne constitue pas un indicateur permettant de les distinguer

La dernière hypothèse envisagée consiste à poser que (i) les deux registres obtenus correspondent bien aux registres souhaités et que (ii) la distinction entre une conversation guidée, c'est-à-dire relevant du registre soutenu, et une conversation libre, c'est-à-dire relevant d'un registre familier, n'est pas significative du point de vue de la liaison.

Dans PFC, la seule différence significative qui est observée s'atteste entre la lecture de documents et la conversation considérée de façon générale. Cette hypothèse est étayée par deux auteurs : (i) Mok (1966) et (ii) Delattre (1947).

(i) Mok (1966), tout d'abord, estime que la distinction entre l'informel et le formel s'amenuise et tend à disparaître :

Au niveau de la conversation familière et de la conversation soignée, certaines liaisons, stables il y a encore quelque temps, deviennent de plus en plus flottantes et tendent à tomber en désuétude, tandis que les liaisons se maintiennent au niveau de la diction et du discours. (Mok, 1966 : 28)

Toutefois, cette hypothèse pose elle aussi un problème. Supposer que la distinction entre les deux registres de parole n'est plus, implique de convenir qu'elle l'a été. Si les résultats des études précédentes (par ex. Martinon, 1913 ; Nyrop, 1934 ; Fouché, 1959 ; Ågren, 1973) sont exacts, on devrait retrouver cette tendance à liasonner plus fréquemment en conversation guidée qu'en conversation libre chez les locuteurs les plus âgés. Ce qui n'est pas le cas, puisque chez les locuteurs PFC âgés de plus de 60 ans, la conversation guidée et la conversation libre ne présentent une différence de fréquence de liaison que de 1,7 % (51,7 % guidée et 50 % libre).

Cependant, comme nous l'avons signalé dans le premier chapitre de ce travail, la plupart des études citées portent sur une population particulière. Les normativistes (Martinon, 1913 ; Grammont, 1914 ; Fouché, 1959 par exemple) aussi bien que les descriptivistes (Nyrop, 1934 ; Delattre, 1966) n'étudient les réalisations de la liaison que chez les locuteurs les plus cultivés. Les études empiriques les plus poussées et les plus récentes ne présentent que les réalisations de la liaison dans des contextes formels : Ågren (1973) étudie la liaison dans les parlers radiophoniques et Encrevé (1988) dans les discours des hommes politiques.

(ii) Le second auteur auquel nous faisons référence est Delattre (1947). Sa description quant aux possibilités de réalisations de la liaison selon le style employé nous semble éclairante :

Dans la conversation familière, on ne fait pas ou presque pas de liaisons facultatives, [...] dans la conversation soignée, on en fait une petite proportion. (Delattre, 1947 : 149)

Selon ses propres dires, il n'y aurait donc aucune différence véritablement significative de la fréquence de la liaison variable entre les deux types de conversation. L'opposition de formalité ne concerne pas les deux types de conversations mais bien la conversation de façon générale et la conférence, situation bien particulière pour laquelle il ajoute qu' « *on en fait [des liaisons variables] beaucoup* ». Nos résultats coïncident

avec les conclusions de Delattre : le style de discussion employé n'a qu'un très faible impact sur la fréquence de la liaison.

Finalement, une fois la première hypothèse éliminée, rien ne permet de trancher entre les deux autres hypothèses. Si l'on considère la seconde hypothèse comme exacte, il est possible que les deux registres obtenus (standard et familier) ne présentent aucune distinction dans la réalisation de la liaison. Si l'on considère la troisième hypothèse, il est possible que la différence de production de la liaison entre une conversation formelle et une conversation informelle ne soit pas significative. La seule vraie distinction possible porterait sur lecture et conversation de façon générale.

6.2.4.2 Variation diastratique

Nous n'avons pas pu confirmer l'importance de la dimension diastratique. En effet, sur les deux paramètres testés (niveau d'études et tranche d'âge), seul l'âge semble avoir une influence déterminante sur la fréquence de la liaison.

Pour l'ensemble des contextes déterminés de liaison, nous avons établi que plus le locuteur était âgé, plus la fréquence de la liaison était élevée. Ces résultats s'inscrivent dans la lignée de ceux de Malécot (1975), Ashby (1981) et Ranson (2008) présentés dans le chapitre 3 (cf. 3.3.3.3 : 92) et confirment l'importance de l'âge dans la réalisation de la liaison variable.

En revanche, le niveau d'études ne semble avoir aucune influence sur la fréquence de la réalisation.

Laks (2005b : 114) rappelle que l'écrit a un rôle prépondérant sur la réalisation de la liaison :

L'écrit se généralise [...] et impose sa légitimité jusqu'à contraindre les usages oraux. [...] la scolarité obligatoire couvre au minimum dix à douze des premières années de la vie et [...] l'apprentissage explicite de l'orthographe y est central. Quotidiennement exercé et renforcé, l'enseignement de la grammaire scolaire, avec ses manipulations de listes de mots, de règles et d'exceptions et ses raisonnements paradigmatiques exemplaires, est au cœur du dispositif pédagogique.

Dans une étude effectuée auprès d'une population du Val Fourré et de ses environs (Mantes la Jolie, 78)¹¹⁰, un certain nombre de locuteurs ayant effectué des études supérieures poussées (Bac+3/ Bac+4) ont un comportement linguistique similaire (*cf.* Mallet, 2001). Ils présentent un taux de réalisation de la liaison plus faible que la plupart des autres locuteurs dont le niveau d'études est bien moindre. Ce faible taux de réalisation de la liaison s'explique en partie par les emplois répétés de structures de type impersonnel (*c'est*+).

De façon plus générale, nous en sommes arrivée à constater que les proportions de réalisation de liaisons variables sont (i) différentes selon chaque locuteur, (ii) sont indépendantes du niveau d'études et (iii) ne peuvent finalement s'expliquer que par une étude au cas par cas (*cf.* Mallet, 2001 : 99-101) :

Cette variation ne dépend apparemment pas du niveau d'études du locuteur, du moins en conversation familière. Il se trouve, en effet, que le pourcentage minimal de réalisation est de 0 %. On relève ce taux parmi des locuteurs de différents niveaux : niveau 1, niveau 2, niveau 3, mais encore plus étonnant, chez des locuteurs du niveau 4¹¹¹. [...] Ces résultats ne nous permettent pas de vérifier l'hypothèse selon laquelle les locuteurs les plus scolarisés lient plus facilement. (Mallet, 2001 : 94-95)

On pourrait alors présumer que l'influence de l'écrit, décrit comme déterminant dans l'apprentissage de la liaison, est combinée à d'autres facteurs beaucoup plus complexes tels que l'utilisation de la variation stylistique. Plus un locuteur est scolarisé, plus il a tendance à utiliser des constructions pour lesquelles il ne réalise la liaison que de façon variable.

Cependant, cet argument n'explique en aucune façon pourquoi les Bac+2 à Bac+4 réalisent plus de liaisons que les Bac à Bac+2, ni pourquoi l'âge semble aussi déterminant dans la fréquence de la liaison. Selon Encrevé (1988 : 259), il faudrait procéder à une analyse très fine de chaque locuteur, mêlant l'ensemble des paramètres stylistiques et sociolinguistiques (âge, sexe, niveau d'études, catégorie socioprofessionnelle, expérience personnelle, *etc.*).

¹¹⁰ L'étude en question ne porte que sur un échantillon de 10 locuteurs. Toutefois, une analyse plus fine des données, locuteur par locuteur, permet de faire émerger la variation inhérente que peut présenter chaque individu.

¹¹¹ Les niveaux croissent à mesure que le niveau d'études s'élève. Un niveau 2 correspond à un niveau scolaire supérieur à un niveau 1, un niveau 3 supérieur à un niveau 2, *etc.*

Par ailleurs, le parcours de vie de locuteurs appartenant à la même génération peut en même temps être différent à un point tel et en même temps commun sur un certain nombre de principes d'éducation, *etc.* qu'il devient véritablement compliqué d'extraire chacun de ces paramètres pour en établir une moyenne. Cette entreprise, outre sa complexité, n'est de toute façon pas garante d'un résultat significatif ou explicatif :

[...] pour en rendre compte globalement, il serait nécessaire d'élaborer un modèle très complexe articulant systématiquement toute une série d'informations sociales qu'il faudrait réunir par une enquête sociologique spécifique. S'agissant d'un fait linguistique explicitement « manipulé » par le système d'enseignement comme l'est la liaison, il faudrait certainement, outre l'origine sociale et le capital scolaire, prendre en compte la notion de génération, entendue principalement au sens de mode de génération, c'est-à-dire non comme classe d'âge biologique mais historique. Une génération réunit des agents ayant participé (différemment) à un même état du système d'enseignement (défini par le mode de recrutement et de formation des enseignants, par le mode d'accès des enseignés à tel niveau d'enseignement, par la durée de la scolarité obligatoire, *etc.*) (Encrevé, 1988 : 259)

Le locuteur présente de toute façon de la variation inhérente. Cette variation résulte d'une multitude de paramètres situationnels non contrôlables :

[...] à l'intérieur d'une même prise de parole publique nous avons observé une dispersion frappante des liaisons facultatives réalisées, et plus encore des liaisons non-enchaînées (qui apparaissent le plus souvent en séries dont la localisation dans le temps de la prise de parole est variable et semble aléatoire). C'est dire que le degré de tension (de surveillance, de formalité) varie (dans certaines limites) à l'intérieur d'un même style de parole sur un marché unique. Il est probable que le contenu du discours contribue à cette variation, mais aussi, dans une intervention publique de longue durée (conférence de presse par exemple), des facteurs tels que la mise en train, la fatigue, l'estimation que fait le locuteur de la réception du discours de l'auditoire, *etc.* : dans ces cas, nous avons constaté de fortes variations de proportion entre début et fin d'une part, partie intermédiaire de l'autre. (Encrevé, 1988 : 258)

Dans le cas des enquêtes PFC, on pourrait ajouter aux paramètres proposés par Encrevé (1988) pour les discours d'hommes politiques, le stress généré par l'appareil enregistreur de la conversation mais également la présence de plusieurs personnes enquêtrices. Ces spéculations ne sont cependant pas vérifiables. Si les enquêteurs PFC ont effectivement signalé les conditions dans lesquelles chacun des entretiens s'est

déroulé, il s'avère au final extrêmement difficile d'établir avec précision ce qui a pu ou non influencer le locuteur.

Dans la section suivante, nous revenons sur l'analyse d'un certain nombre de ces paramètres et considérons les conversations libre et guidée. On utilisera la lecture de texte en complément des analyses quand cela sera nécessaire.

6.3 Les consonnes de liaison

Dans cette section, nous proposons de passer en revue les différentes consonnes de liaison qui apparaissent dans le corpus PFC que nous avons établi, et d'examiner leur fréquence d'apparition.

La liaison est avant tout un phénomène phonologique qui met en jeu l'alternance consonne / zéro. Il apparaît donc important de tester les différents niveaux d'influence que peuvent exercer les contraintes phonologiques que nous avons décrites précédemment (*cf.* 3.1), comme l'influence de la nature de la CL sur la réalisation de la liaison (*cf.* aussi Pagliano et Laks, 2005 : 5).

Indépendamment de tout autre facteur, les consonnes de liaison ne sont généralement pas considérées comme un paramètre explicite d'influence de la réalisation de la liaison. Leur analyse est le plus souvent corrélée à des analyses syntaxique et lexicale (*cf.* Ågren, 1973 ; Malécot, 1975 ; Pagliano et Laks, 2005 ; Durand et Lyche, 2005, entre autres). Toutefois, un certain nombre d'hypothèses concernant leur nature et leur fréquence d'apparition ont été présentées :

1. les CL /p, k, r/ sont très peu représentées en contexte de liaison et ne lient pratiquement jamais ;
2. au contraire, les CL /n, t, z/ sont très fréquentes : elles sont très représentées en contexte de liaison et elles lient souvent ;

3. la fréquence de distribution est différente selon les travaux. Pour certains, on aurait une liaison réalisée de façon décroissante en /z/, /n/ et /t/ (proposition 1.). Pour d'autres, on aurait une liaison plus réalisée en /z/ qu'en /t/ et en /n/ (proposition 2.). Enfin, pour les derniers, la liaison serait réalisée de façon décroissante en /n/, /z/ et /t/ (proposition 3.). Ces trois principales propositions sont résumées dans le tableau ci-dessous :

Proposition 1 [z] > [n] > [t]	Proposition 2 [z] > [t] ≥ [n]	Proposition 3 [n] > [z] > [t]
Green et Hintze (2001 : 34)	Léon (1992 : 152)	Malécot (1975 : 164)
Pagliano et Laks (2005 : 5)		
Durand et Lyche (2008 : 58)		

Figure 6-13 Les propositions de distribution des CL /n/, /t/ et z/

Une analyse des CL en termes de distribution, en termes d'ordre d'apparition dans la liaison, amène à se questionner sur la signification d'une affirmation du type « X liaisonne plus que Y ». Si par exemple, la CL /z/ manifeste plus de liaisons que la CL /n/, peut-on considérer que les résultats obtenus s'appuient sur les mêmes faits ? Peut-être que le nombre de « mots » liaisonnant en /z/ est très restreint et correspond à des contextes dans lesquels la liaison est toujours réalisée, alors que le nombre de « mots » liaisonnant contenant la CL /n/ sont peut-être plus nombreux et se partagent entre liaisons toujours réalisées, fréquemment réalisées, peu réalisées et jamais réalisées. Sachant que la liaison est, comme la définissent Pagliano et Laks (2005 : 7), « [...] un phénomène multimodal entremêlant écrit et oral, représentation en mots et représentation en groupes de souffle, fréquences d'usage et variations sociostylistiques, réinterprétations fonctionnelles morpho-syntaxiques et cohésion sémantique [...] », quel est l'intérêt d'une analyse de la nature de la CL et de son taux de réalisation de la liaison ? Ou plutôt, peut-on se satisfaire d'une telle analyse pour tenter d'expliquer le fonctionnement de la liaison, même en partie ? Bien évidemment, non. Tous les travaux sur la question, notamment ceux auxquels nous faisons référence et qui proposent un ordre dans la distribution des CL, ne se contentent pas d'observer et de calculer le nombre de contextes réalisant la liaison et ceux ne réalisant pas la liaison. Ils analysent leurs résultats en fonction d'autres dimensions (lexicales, syntaxiques, etc.). Pagliano et Laks (2005), par exemple, de la même façon que Martinon (1913), Morin et Kaye

(1982) et Durand et Lyche (2008), expliquent la très forte fréquence de la liaison en /z/ par le biais d'une analyse morpho-syntaxique (*cf.* également 3.2.1, pp. 59 et *sq.*) :

[Les consonnes de liaison] sont au nombre d'une dizaine, seules 4 sont présentes dans notre corpus ([z, n, t, r]), deux autres étant un peu moins fréquentes ([p, k]), soit six consonnes sur plus de dix consonnes graphiques possibles. De plus, en ce qui concerne l'enquête lyonnaise, seules les coronales sont représentées, avec une prépondérance de [z], marquant ainsi l'importance de la liaison de pluriel ainsi réalisée dans le total des liaisons effectuées. (Pagliano et Laks, 2005 : 5)

Aussi, la mise en évidence des CL qui apparaissent le plus dans des contextes de liaison permet :

1. de vérifier si leur (non)-réalisation est corrélée à leur « potentiel » de production ;
2. de mettre au jour les unités lexicales en jeu dans la liaison et valider différentes hypothèses mises en avant par un certain nombre de travaux : le rôle morpho-phonologique que peut avoir la CL (par ex. /z/ dans un contexte pluriel) ; le rôle du contexte syntaxique dans lequel est pris le « mot » contenant la CL, *etc.* ;
3. une fois les unités lexicales isolées, de procéder à une analyse lexicale du phénomène : tester la longueur des éléments et la complexité phonologique des éléments en jeu ;
4. de façon beaucoup plus générale, de poser un certain nombre de problèmes relatifs aux autres dimensions en jeu dans le phénomène.

L'analyse de l'influence de ces autres paramètres, lexicaux notamment, ne peut se faire que si l'on considère les fréquences de réalisation des CL de façon très générale.

Aussi, (i) au vu du nombre de contextes dont nous disposons pour identifier les contextes dans lesquels chacune des CL apparaît ; (ii) au vu des métadonnées dont nous disposons (*cf.* le codage liaison et ses limites, 5.1.3.3) ; et (iii) étant donné notre objectif, nous avons eu recours à une identification par liste de mots (mots fréquents, mots analysés dans des travaux, *etc.*).

Cette démarche influence nécessairement la présentation de nos résultats. Il importe de conserver en mémoire que notre objectif, ici, est uniquement de présenter les résultats en terme de fréquence de CL bien que, *de facto*, nous fournissions des

informations qui relèvent d'autres dimensions d'analyse (notamment syntaxique et lexicale). Un même mot peut en effet présenter une liaison variable ou catégorique selon le contexte dans lequel il est employé. Prenons le cas de *un* : les exemples (70)¹¹² et (71)¹¹³ correspondent respectivement à un contexte de liaison réalisée et à un contexte de liaison non réalisée.

Or, il se trouve que dans le premier cas, *un* est employé comme déterminant, alors que dans le second cas, il est pronom¹¹⁴. Ces deux exemples, d'un point de vue syntaxique, ne devraient pas être traités ensemble mais être clairement distingués ; ce que nous ne faisons pas dans cette partie, puisque étudiant uniquement la réalisation de la CL, nous les mettons sur le même plan. Ceci implique nécessairement qu'à ce stade, les informations fournies seront incomplètes.

Certes, on ne peut se satisfaire d'une simple analyse en termes de statistiques de la réalisation des CL : Martinon (1913), dans son traité de phonétique, fournit une description des CL qu'il relie toujours à la construction syntaxique (par ex. les labiales ne lient pas sauf *p* dans les adverbes *beaucoup et trop* devant participe ou devant la préposition *à*). Concernant la CL /z/, Gougenheim (1935), met en avant sa valeur morpho-phonologique même après *r* et compare *un cours élémentaire*, sans /z/ et *des cours élémentaires*, avec /z/ (cf. 3.2.1). Mais cette présentation nous permet de mettre en avant une série de problèmes que nous proposons en bilan de cette section et qui seront traités ultérieurement. Elle a aussi l'avantage de mettre en avant à quel point les différentes dimensions en jeu dans le fonctionnement de la liaison sont interdépendantes.

Pour ce faire, nous procédons en deux étapes :

1. Dans un premier temps, nous faisons l'inventaire de l'ensemble des consonnes de liaison et étudions leur distribution (cf. 6.3.1). Cela revient dans l'ensemble du corpus :

¹¹² *sur un11n an.*

¹¹³ *chez l'un10 ou l'autre.*

¹¹⁴ Nous verrons en 6.5 que ce problème, combiné au nombre très important que peuvent présenter certaines formes, nous a limité dans le nombre de catégories syntaxiques traitées.

- à distinguer le sous-ensemble des contextes de liaisons réalisées et le sous-ensemble des contextes de liaisons potentielles non réalisées ; par la suite, pour simplifier les formulations, nous désignerons le sous-ensemble des contextes de liaisons réalisées par **ELR** et le sous-ensemble des contextes de liaisons non réalisées par **ELNR** ;

- ensuite, à compter dans ELR le nombre d'occurrences de chaque CL.

2. Dans un second temps, pour chaque CL, nous construisons un corpus d'observation (donc autant de corpus d'observation que de CL), entendu comme l'ensemble des contextes où la liaison est réalisée et l'ensemble des contextes où elle n'est pas réalisée.

Autrement dit, nous construisons les ensembles suivants :

Corpus d'observation	Contenu
<i>Ek</i>	<p>Comprend les deux ensembles :</p> <ul style="list-style-type: none"> - ELRk (ensemble des contextes de liaisons réalisées avec les CL /g/ et /k/); - et ELNRk (ensemble des contextes où la liaison avec les CL /g/ ou /k/ n'est pas réalisée).
<i>En</i>	<p>Comprend les deux ensembles :</p> <ul style="list-style-type: none"> - ELRn (ensemble des contextes de liaisons réalisées avec la CL /n/); - et ELNRn (ensemble des contextes où la liaison avec la CL /n/ n'est pas réalisée).
<i>Ep</i>	<p>Comprend les deux ensembles :</p> <ul style="list-style-type: none"> - ELRp (ensemble des contextes de liaisons réalisées avec la CL /p/); - et ELNRp (ensemble des contextes où la liaison avec la CL /p/ n'est pas réalisée).
<i>Er</i>	<p>Comprend les deux ensembles :</p> <ul style="list-style-type: none"> - ELRr (ensemble des contextes de liaisons réalisées avec la CL /r/); - et ELNRr (ensemble des contextes où la liaison avec la CL /r/ n'est pas réalisée).
<i>Et</i>	<p>Comprend les deux ensembles :</p> <ul style="list-style-type: none"> - ELRt (ensemble des contextes de liaisons réalisées avec la CL /t/); - et ELNRt (ensemble des contextes où la liaison avec la CL /t/ n'est pas réalisée).
<i>Ez</i>	<p>Comprend les deux ensembles :</p> <ul style="list-style-type: none"> - ELRz (ensemble des contextes de liaisons réalisées avec la CL /z/); - et ELNRz (ensemble des contextes où la liaison avec la CL /z/ n'est pas réalisée).

Figure 6-14 Ensemble des corpus d'observation créés pour chaque CL

Pour parler de manière générale de l'ensemble des corpus construits à partir de chaque CL, nous utiliserons l'abréviation **ECL** (soit *Ez + Et + En + etc.*).

Cela nous permet, d'une part d'identifier le taux de réalisation de chaque CL en fonction de tous les contextes possibles du corpus (*cf.* 6.3.2, pp. 212 et *sq.*) ; et d'autre part, de déterminer si l'opposition entre les CL /p, k, r/ et les CL /n, t, z/ est pertinente.

Une fois cette opposition vérifiée, cela nous permet d'étudier précisément, pour chaque CL, quels sont les contextes dans lesquels une CL donnée apparaît et ceux dans lesquels elle ne se réalise pas (*cf.* 6.3.3, pp. 214 et *sq.*), tout en constituant deux groupes : (i) le groupe des CL qui ne seront pas considérées dans la suite de nos analyses (/p, k, r/) (*cf.* 6.3.3.1) ; et (ii) le groupe des CL qui seront considérées (/n, t, z/) (*cf.* 6.3.3.2).

Les abréviations utilisées tout au long de cette section sont reprises et résumées dans le tableau ci-dessous :

Abréviation	Définition
ECL	Ensemble des corpus ($Ek + En + etc.$) construits pour chaque CL.
E $\left(\begin{array}{c} g \\ k \\ n \\ p \\ t \\ r \\ z \end{array} \right)$	Corpus d'observation contenant l'ensemble des contextes ELR et ELNR de chaque CL.
ELR	Sous-ensemble des contextes de liaisons réalisées.
ELNR	Sous-ensemble des contextes de liaisons non réalisées.

Figure 6-15 Tableau des abréviations utilisées pour l'analyse des CL

6.3.1 Inventaire et distribution des consonnes de liaison

Lorsqu'une consonne de liaison n'est pas réalisée, le moteur de recherche PFC ne signale pas quelle consonne est « absente ». Il indique uniquement que la liaison n'a pas eu lieu, sans présumer de la consonne hypothétiquement réalisable. Cette démarche permet d'éviter de prédire la production d'une consonne orale sur la base d'une représentation graphique.

Cependant, pour les besoins de la discussion, nous admettons ici que la production orale, si elle avait été réalisée, correspondrait à la consonne graphique¹¹⁵.

Aussi, pour calculer les contextes où la consonne de liaison n'est pas produite, nous avons cherché, pour chaque consonne codée, les contextes où :

- il y a absence de liaison (codage 10 ou 20) ;
- le contexte droit (M2) se termine par la consonne orthographique {d, g, k, n, p, r, s, t, x et / ou z} ;
- le contexte gauche (M1) commence par la voyelle orthographique {a, e, i, o, u, y et / ou « h aspiré »}.

Nous donnons ici la distribution des consonnes de liaison sur l'ensemble des liaisons réalisées dans les conversations libre et guidée, soit un total de 12 841 liaisons réalisées :

- les consonnes /n/, /t/ et /z/ sont les trois CL les plus fréquentes en liaison dans le corpus considéré. Elles représentent respectivement 46,4 %, 38,7 % et 14,7 % des liaisons réalisées.
- les consonnes /r/, /p/, /g/ et /k/ (dorénavant réunies en « autres ») représentent 0,2 % des liaisons réalisées.

La Figure 6-16 schématise cette distribution :

¹¹⁵ Les liaisons réalisées avec une consonne de liaison différente de la CL attendue (*il a dépensé cent euros* réalisé [iladepãsesãzœro] au lieu de [iladepãsesãtœro]) ne sont pas prises en compte ici. Rappelons qu'elles font partie des contextes que nous avons éliminés (cf. 6.1.1, pp. 179 et sq.).

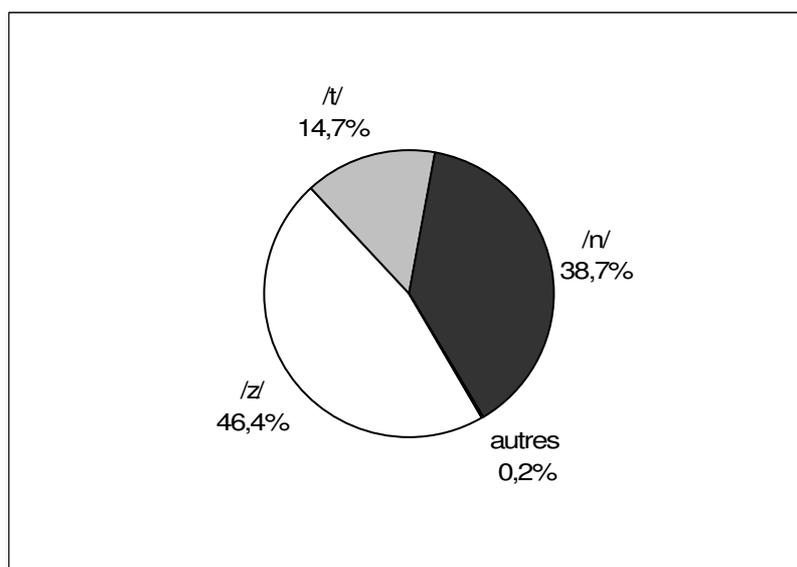


Figure 6-16 Distribution des CL dans le corpus comprenant les liaisons réalisées

Maintenant que nous avons établi quelles étaient les consonnes de liaison les plus fréquentes dans ELR, à savoir /z/, /n/ et /t/, observons quelle est la fréquence de réalisation de chacune de ces consonnes dans son propre ECL.

6.3.2 Fréquence de chaque CL dans chaque ECL

Nous examinons maintenant quelle est la fréquence de chaque CL dans son ECL. En effet, qu'une CL ait un taux de réalisation élevé dans ELR ne signifie pas pour autant qu'elle soit la consonne qui liaisonne le plus dans son ECL, comparativement aux autres CL envisagées chacune dans leur propre ensemble (ECL).

Après avoir dénombré les contextes potentiels d'apparition de chacune des consonnes de liaison présentées en situation conversationnelle (guidée et libre), soit l'identification du cardinal de chaque ECL, nous avons calculé les taux de réalisation et de non-réalisation pour chacune d'entre elles. Dans la Figure 6-17, qui présente ces calculs, nous avons classé les consonnes selon le cardinal de chaque ECL.

CL	ECL	Liaisons réalisées		Liaisons non réalisées	
		Nombre total	% de L	Nombre total	% de L
/z/	13 852	5 960	43 %	7 892	57 %
/t/	8 113	1 888	23,3 %	6 225	76,7 %
/n/	5 512	4 970	90,2 %	542	9,8 %
/r/	942	12	1,3 %	930	98,7 %
/p/	98	11	11,2 %	87	88,8 %
/k/	2	0	0 %	2	100 %
Total	28 519	12 841	45 %	15 678	54 %

Figure 6-17 Fréquence des liaisons réalisées et non réalisées dans chaque ECL

Les résultats indiquent que /z/ arrive en tête et comptabilise le plus grand nombre de contextes possibles (soit 13 852) ; en dernier, nous avons /k/ qui comptabilise le nombre de contextes potentiels le plus faible (soit 2).

Ces données sont présentées sous forme graphique dans la Figure 6-18 :

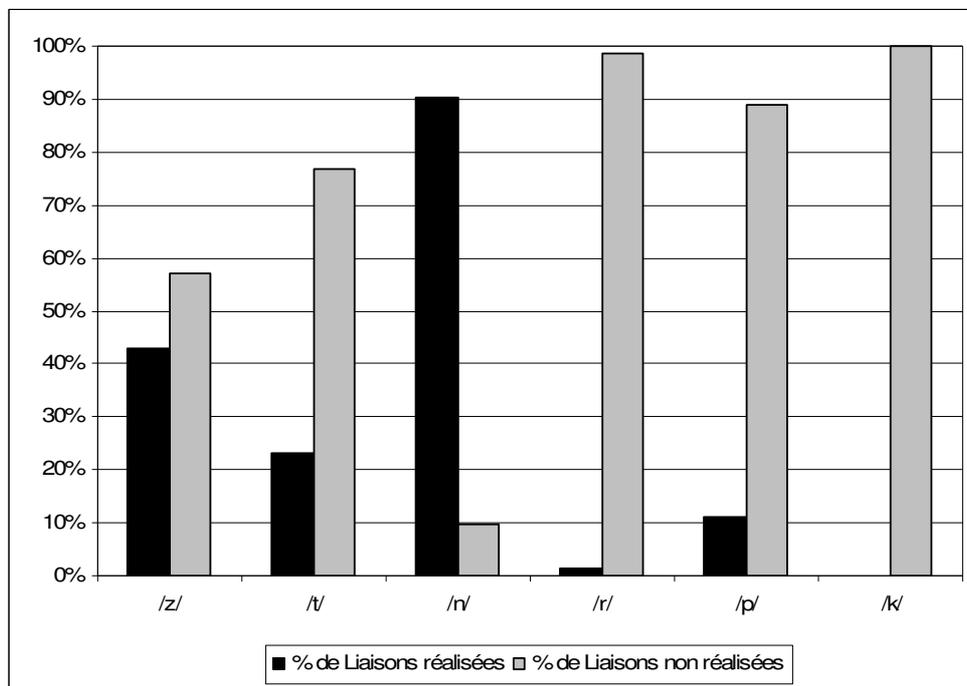


Figure 6-18 Fréquence des liaisons réalisées et non réalisées dans chaque ECL sous forme graphique

Ceci nous amène à constater que le nombre de liaisons réalisées n'est pas proportionnel au nombre de contextes potentiels. Par exemple, les contextes de liaisons possibles relevés pour la consonne /z/ sont les plus nombreux, 13 852 au total, et le taux de réalisation de la liaison s'élève à 43 % dans *Ez*. Comparativement, le taux de réalisation de la consonne de liaison /n/, susceptible de se réaliser dans 5 512 sites (soit plus de deux fois moins que le nombre de sites associés à /z/) s'élève à 90,2 % dans *En*. On remarque également que le taux de réalisation de la consonne /t/ ne s'élève qu'à 23,3 % pour un nombre total de 8 113 sites, alors que le taux de réalisation de /n/ est de 90,2 % pour un nombre total de 5 512.

Les différences dans les cardinaux de chaque ECL, de même que les taux de réalisation / non-réalisation de la liaison demandent de s'interroger plus précisément sur les contextes que nous avons considérés.

6.3.3 Analyse contextuelle d'apparition des consonnes de liaison

L'analyse contextuelle d'apparition des consonnes de liaison ne porte pas sur la totalité des CL que nous venons de présenter. Par ailleurs, concernant les consonnes de liaison considérées, nous ne fournissons qu'un sous-ensemble des contextes.

6.3.3.1 Les consonnes de liaison non considérées

Les consonnes de liaison /r/, /p/ et /k/ ne seront pas examinées en détail. Cependant, nous en donnons ici les principes de fonctionnement généraux.

Nous reprenons dans le tableau ci-dessous l'ensemble de ces trois cas que nous excluons par la suite :

Consonne de liaison	Nombre de contextes de liaisons	Nombre de LR	Nombre de LNR
/r/	942	12	930
/p/	98	11	87
/k/	2	0	2
Total	1 042	23	1 019

Figure 6-19 Liaisons réalisées et non réalisées dans les ECL non considérés (*Er*, *Ep* et *Ek*)

6.3.3.1.1 La consonne /r/

Sur un total de 942 contextes possibles de liaison, on ne compte que 12 réalisations (cf. Figure 6-20) ; ce qui ne représente que 1,3 % de ses possibles.

	<i>Er</i>	ELRr	ELNRr
Nombre	942	12	930
%	100 %	1,3 %	98,7 %

Figure 6-20 Fréquence des liaisons réalisées et non réalisées dans *Er*

Après examen du corpus, il s'avère que la consonne de liaison /r/ se concentre particulièrement à la finale des verbes infinitif du premier groupe (par ex. : *donner*), lesquels ne réalisent pratiquement jamais la liaison et intègrent la catégorie des liaisons variables. La plupart de ces contextes, Verbe +, comporte un verbe bi- ou trisyllabique, comme par exemple :

- (53) je voulais rester₂₀ à la propriété (11aall) : le verbe *rester* est bisyllabique ;
- (54) je voudrais bien retourner₂₀ à l'école (91aell) : le verbe *retourner* est trisyllabique.

6.3.3.1.2 La consonne /p/

Sur un total de 98 contextes possibles de liaison en /p/, on ne compte que 11 réalisations (cf. Figure 6-21).

	<i>Ep</i>	<i>ELRp</i>	<i>ELNRp</i>
Nombre	98	11	87
%	100 %	11,2 %	88,8 %

Figure 6-21 **Fréquence des liaisons réalisées et non réalisées dans *Ep***

On ne dénombre que deux occurrences pouvant liasonner en /p/ sur l'ensemble du corpus. Ce sont des termes invariables courants manifestant une liaison variable (cf. Figure 3-14 : 80) :

1. *trop* monosyllabique présent 52 fois en contexte de liaison et qui réalise 8 liaisons dont nous donnons deux exemples en (55) et en (56). Des exemples de liaisons non réalisées sont présentés en (57) et (58) :

(55) *pas/ encore trop11p évident (13aag1)*

(56) *que je sois trop11p âgé (svajb1)*

(57) *trop10 âgée euh (21amb1)*

(58) *un peu trop10 elle-même (81aaa1)*

2. et *beaucoup* bisyllabique, présent dans 46 contextes de liaison et qui réalise 3 liaisons (exemples (59), (60) et (61)). Des exemples de liaisons non réalisées sont présentés en (62) et (63) :

(59) *ils se livrent beaucoup21p à nous (64ama2)*

(60) *beaucoup21p occupée (54bpm1)*

(61) *qui m'a beaucoup21p impressionné (ciafk1)*

(62) *les enfants m'ont beaucoup20 aidé (13bjc1)*

(63) *a beaucoup20 évolué dans sa forme (69ajl1)*

Les taux de réalisation de chacune de ces deux occurrences s'élèvent respectivement à environ 15,4 % (*trop*) et à un peu plus de 6,5 % (*beaucoup*).

6.3.3.1.3 La consonne /k/

Les contextes potentiels de liaison en /k/ sont les plus rares. Sur la totalité des sites potentiels (soit 28 519), on ne dénombre que deux occurrences, pour lesquelles la liaison n'est pas réalisée (cf. Figure ci-dessous) :

	<i>Ek</i>	ELRk	ELNRk
Nombre	2	0	2
%	100 %	0 %	100 %

Figure 6-22 **Fréquence des liaisons réalisées et non réalisées dans *Ek***

Ces deux occurrences n'ont pas la même consonne graphique finale : la première contient la consonne orthographique {g} (exemple (64)) et la seconde la consonne orthographique {c} (exemple (65)) :

(64) *un peu long10 à vous déclarer (svajb1)*

(65) *pour être franc10 avec vous (abahs1)*

Les consonnes de liaison /r/, /p/ et /k/ ne seront pas exploitées dans les sections suivantes. Nous les considérons comme intégrant des contextes de liaisons marginaux (« autres ») du fait de leur très faible représentation sur l'ensemble des liaisons réalisées (soit 0,2 %) et de leur très faible représentation sur l'ensemble des contextes potentiels de liaison.

6.3.3.2 Les consonnes de liaison considérées

Nous limitons l'analyse contextuelle aux trois consonnes qui manifestent le plus de variation : /n/, /t/ et /z/. Le nombre total de contextes considérés s'élève désormais à 27 477.

	Nombre de contextes codés
Liaisons en /r/, /p/ et /k/	1 042
Liaisons en /n/, /t/ et /z/	27 477
Nombre total de contextes dans E	28 519

Figure 6-23 Consonnes de liaison considérées

Nous reprenons dans le tableau ci-dessous l'ensemble des contextes de ces trois cas que nous allons observer :

Consonne de liaison	Nombre de contextes de liaisons	Nombre de LR	Nombre de LNR
/n/	5 512	4 970	542
/t/	8 113	1 888	6 225
/z/	13 852	5 960	7 892
Total	27 477	12 818	14 659

Figure 6-24 Liaisons réalisées et non réalisées dans les ECL considérées (*En, Et et Ez*)

L'ensemble des contextes en /n/, /t/ et /z/ n'ayant pu être identifié, les résultats présentés ne portent que sur un sous-ensemble des contextes présents dans le corpus. Afin d'être le plus clair possible, nous avons de nouveau recours à un système d'abréviations :

1. EI : corpus d'observation contenant l'ensemble des contextes contenus dans EILR et les EILNR ;
2. EILR : sous-ensemble des contextes de liaison identifiés ;
3. et EILNR : sous-ensemble des contextes de non-liaison identifiés.

Les abréviations utilisées sont reprises et résumées dans le tableau ci-dessous :

Abréviations	Définition
EI	Corpus d'observation contenant l'ensemble des contextes ELR et ELNR identifiés pour l'ensemble des consonnes /n/, /t/ et /z/
EILR	Sous-ensemble des contextes de liaison identifiés
EILNR	Sous-ensemble des contextes de non-liaison identifiés

Figure 6-25 Tableau des abréviations utilisées pour l'analyse contextuelle des consonnes /n/, /t/ et /z/

Sur l'ensemble du corpus d'observation E, le nombre total de contextes potentiels de liaison en /n/, /t/ et /z/ s'élève à 27 477. Nous avons identifié 17 775 de ces contextes, soient 64,7 %.

On remarque que ces identifications sont inégales. Il se trouve en effet que le nombre de contextes de liaisons réalisées identifiés (EILR) est nettement supérieur au nombre de contextes de liaisons non réalisées identifiés (EILNR) et ce, pour les trois CL.

De façon générale, 84,4 % des EILR et 47,4 % des EILNR sont désormais connus.

Les résultats pour chaque consonne considérée sont présentés dans le tableau suivant :

	EI dans ECL		EILR dans ELR		EILNR dans ELNR	
	Nombre	%	Nombre	%	Nombre	%
<i>/n/</i>	4 366	79,2 %	4 229	85,1 %	137	25,3 %
<i>/t/</i>	4 972	61,3 %	1 610	85,3 %	3 362	54 %
<i>/z/</i>	8 437	60,9 %	4 984	83,6 %	3 453	43,8 %
Total	17 775	64,7 %	10 823	84,4 %	6 952	47,4 %

Figure 6-26 Contextes identifiés pour les CL /z/, /n/ et /t/

Procédons maintenant à l'analyse de chacune des CL considérées.

6.3.3.2.1 La consonne /n/

Le nombre potentiel de contextes de liaison possibles avec la consonne /n/ est de 5 512, répartis de la façon suivante :

- 4 970 contextes de liaisons réalisées ;
- et 542 contextes de liaisons non réalisées.

La liaison en /n/ est donc réalisée dans 90,2 % des contextes contenus dans *En* :

	<i>En</i>	ELRn	ELNRn
Nombre	5 512	4 970	542
%	100 %	90,2 %	9,8 %

Figure 6-27 Fréquence des liaisons réalisées et non réalisées dans *En*

Sur ces 5 512 contextes, nous en avons identifié 4 366, soit 79,2 % d'entre eux (cf. Figure 6-26) :

- 4 229 contextes de liaisons réalisées ;
- et 137 contextes de liaisons non réalisées.

Autrement dit, sur le nombre total de contextes potentiels (*En*), 85,1 % des liaisons réalisées et 25,3 % des liaisons non réalisées ont été identifiées. Ces résultats sont représentés dans le tableau ci-dessous :

	<i>En</i>	<i>EIn</i>	ELRn	EILRn	ELNRn	EILNRn
Nombre	5 512	4 366	4 970	4 229	542	137
%	100 %	79,2 %	100 %	85,09 %	100 %	25,27 %

Figure 6-28 Contextes identifiés pour la CL /n/ (*En*)

Dorénavant, toutes nos remarques s'appuieront sur ces données qui constituent un nouvel ensemble, *EIn*.

Dans *EIn*, la liaison est réalisée dans 96,9 % des contextes. Seuls 3,1 % des contextes ne réalisent pas la liaison (cf. Figure 6-29) :

	<i>EIn</i>	EILRn	EILNRn
Nombre	4 366	4 229	137
%	100 %	96,86 %	3,13 %

Figure 6-29 Fréquence des liaisons réalisées et non réalisées dans *EIn*

On peut donc dire que, dans *EIn*, les contextes de liaison manifestent presque tous une liaison catégorique.

L'inventaire de ces contextes fait apparaître que le nombre de mots liaisonnants (M1 différents) est faible. On en dénombre 6 répartis de la façon suivante : le pronom personnel *on*, les possessifs *mon* et *son*, *un*, et les « invariables » *en* et *bien*.

Les taux de réalisation de la liaison pour chacune de ces formes sont présentés dans le tableau ci-dessous :

	M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
1	<i>on</i>	2 489	2 489	0	100 %
2	<i>mon</i>	124	124	0	100 %
3	<i>son</i> ¹¹⁶	58	58	0	100 %
4	<i>un</i>	551	540	11	98 %
5	<i>en</i>	959	939	20	97,9 %
6	<i>bien</i> ¹¹⁷	185	79	106	42,7 %
	Total	4 366	4 229	137	96,9 %

Figure 6-30 M1 contenus dans *EIn*

Détailler le nombre de contextes de liaison ainsi que leur taux de réalisation dans la liaison, amène à constater que :

- (i) la liaison est toujours réalisée pour trois d'entre eux (*on*, *mon* et *son*) ;
- (ii) la liaison est très souvent réalisée pour *un* et *en* (respectivement 98 % et 97,9 % de liaisons réalisées) ;
- (iii) enfin, que la liaison est peu réalisée avec *bien* (42,7 % de liaisons réalisées).

Examinons chacune de ces formes présentes en contextes de liaison.

1. *On*

On est présent dans 2 489 contextes de liaisons. Sur l'ensemble de ces contextes, la liaison est toujours réalisée :

¹¹⁶ Dans le corpus, le mot *son* correspond toujours au possessif et jamais au nom singulier.

¹¹⁷ De la même façon, *bien* n'est jamais identifié comme nom singulier mais toujours comme adverbe.

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>on</i>	2 489	2 489	0	100 %

Figure 6-31 Fréquence des liaisons en /n/ avec *on*

De façon plus générale, *on* représente :

- 57 % des contextes de liaisons dans *EIn*, comme par exemple dans :

(66) *onlIn est vite parti*

- et 58,9 % des liaisons réalisées en /n/ (soit *EILRn*) :

(67) *enlIn Angleterre*

Ces résultats sont présentés dans le tableau suivant :

<i>On</i>	dans <i>EIn</i>	dans <i>EILRn</i>	dans <i>EILNRn</i>
Nombre	2 489 / 4 366	2 489 / 4 229	0 / 5 512
%	57 %	58,85 %	0 %

Figure 6-32 Représentation de *on* dans *EIn*, *EILRn* et *EILNRn*

2. *Mon* et *son*

Mon et *son* sont respectivement présents dans 124 et 58 contextes de liaisons. La liaison est toujours réalisée dans les deux contextes :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>mon</i>	124	124	0	100 %
<i>son</i>	58	58	0	100 %

Figure 6-33 Fréquence des liaisons en /n/ avec *mon* et *son*

De façon plus générale, *mon* représente :

- 2,8 % des contextes identifiés de liaisons en /n/ (soit *EIn*) ;
- et 2,9 % des contextes identifiés de liaisons réalisées en /n/ (soit *EILRn*) :

(68) *c'était monlIn année*

Ces résultats sont présentés dans le tableau ci-dessous :

<i>mon</i>	dans <i>EIn</i>	dans <i>EILRn</i>	dans <i>EILNRn</i>
Nombre	124 / 4 366	124 / 4 229	0 / 137
%	2,84 %	2,93 %	0 %

Figure 6-34 Représentation de *mon* dans *EIn*, *EILRn* et *EILNRn*

Et *son* représente :

- 1,3 % des contextes identifiés de liaisons en /n/ (soit *EIn*) ;
- et 1,4 % des contextes identifiés de liaisons réalisées en /n/ (soit *EILRn*) :

(69) *son l In emploi jeune*

Soit :

<i>son</i>	dans <i>EIn</i>	dans <i>EILRn</i>	dans <i>EILNRn</i>
Nombre	58 / 4 366	58 / 4 229	0 / 137
%	1,33 %	1,37 %	0 %

Figure 6-35 Représentation de *son* dans *EIn*, *EILRn* et *EILNRn*

3. *Un*

Un est présent dans 551 contextes de liaison. Seuls 11 contextes n'ont pas manifesté de liaison, comme l'illustre le tableau suivant :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>un</i>	551	540	11	98 %

Figure 6-36 Fréquence des liaisons en /n/ avec *un*

Les contextes de liaison dans lesquels *un* apparaît comme M1 représentent :

- 12,6 % des contextes identifiés de liaisons en /n/ (soit *EIn*) ;
- 12,8 % des contextes identifiés de liaisons réalisées en /n/ (soit *EILRn*) :

(70) *Sur un l In an*

- et 8 % des contextes qui ne réalisent pas la liaison (dans *EILNRn*) :

(71) *chez l'un10 ou l'autre*

(72) *c'est un10 huit*

Soit :

<i>un</i>	dans <i>EIn</i>	dans <i>EILRn</i>	dans <i>EILNRn</i>
Nombre	551 / 4 366	540 / 4 229	11 / 137
%	12,62 %	12,77 %	8 %

Figure 6-37 Représentation de *un* dans *EIn*, *EILRn* et *EILNRn*

4. *En*

En est présent dans 959 contextes et la liaison est réalisée dans 939 d'entre eux (soit 97,9 %), comme on peut le voir dans le tableau suivant :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>En</i>	959	939	20	97,9 %

Figure 6-38 Fréquence des liaisons en /n/ avec *en*

Les contextes de liaison dans lesquels on retrouve *en* en M1, ainsi que les contextes où cette liaison est réalisée ou non représentent donc :

- 22 % des contextes identifiés de liaisons en /n/ (soit *EIn*) ;
- 22,2 % des contextes identifiés de liaisons réalisées en /n/ (soit *EILRn*) :

(73) *Il faut en11n avoir trois*

(74) *en11n arrière*

- et 14,60 % des contextes de liaison non réalisée en /n/ :

(75) *on allait en10 haut*

(76) *manque beaucoup de confiance en10 elle*

Soit :

<i>En</i>	dans <i>EIn</i>	dans <i>EILRn</i>	dans <i>EILNRn</i>
Nombre	959 / 4 366	939 / 4 229	20 / 137
%	21,96 %	22,20 %	14,59 %

Figure 6-39 Fréquence de *en* dans *EIn*, *EILRn* et *EILNRn*

5. *Bien*

Bien se retrouve dans 185 contextes qui réalisent la liaison à hauteur de 42,7 % (soit 79 liaisons réalisées) :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>bien</i>	185	79	106	42,7 %

Figure 6-40 Fréquence des liaisons en /n/ avec *bien*

Les contextes de liaisons possibles, réalisées et non réalisées représentent :

- 4,2 % des contextes identifiés de liaisons en /n/ (soit *EIn*) ;
- 1,9 % des contextes identifiés de liaisons réalisées en /n/ (soit *EILRn*) :

(77) *s'est bien l In amélioré*

- et 77,4 % des contextes dans lesquels la CL /n/ n'est pas réalisée :

(78) *très bien l0 au cinéma*

Ces résultats sont repris dans le tableau suivant :

<i>bien</i>	dans <i>EIn</i>	dans <i>EILRn</i>	dans <i>EILNRn</i>
Nombre	185 / 4 366	79 / 4 229	106 / 137
%	4,24 %	1,87 %	77,37 %

Figure 6-41 Représentation de *bien* dans *EIn*, *EILRn* et *EILNRn*

Dans notre corpus, la consonne de liaison /n/ n'est présente que dans très peu de mots : 6 au total. La liaison est fréquente, surtout avec *bien*.

6. Bilan

La liaison avec la CL /n/ est réalisée dans la plupart des contextes. Seules 6 formes ont été recensées, ce qui représente près de 80 % de l'ensemble des contextes de liaison en /n/ (cf. Figure 6-27). Trois de ces formes ne lient pas systématiquement : *bien* qui représente 77,4 % des contextes de EILNRn, *en* (14,6 % de EILNRn) et *un* (8 % de EILNRn) :

/n/	<i>EIn</i>	EILRn	EILNRn
<i>on</i>	57 %	58,9 %	0 %
<i>mon, son</i>	4,2 %	4,3 %	0 %
<i>un</i>	12,6 %	12,8 %	8 %
<i>en</i>	22 %	22,2 %	14,6 %
<i>bien</i>	4,2 %	1,9 %	77,4 %
Total	100 %	100 %	100 %

Figure 6-42 Fréquence de réalisation de la liaison des mots dans *EIn*

On remarque d'une part, que les ensembles *EIn* et EILRn sont surreprésentés par *on* qui occupe :

- 57 % de l'ensemble *EIn* ;
- et 58,9 % de l'ensemble EILRn.

Aussi, comparativement aux possessifs *mon* et *son*, dont le nombre de contextes est beaucoup plus restreint (respectivement environ 2,8 % et 1,3 % des contextes de *EIn* et des contextes de EILRn ; cf. Figure 6-34 et Figure 6-35), cette surreprésentation favorise-t-elle probablement un taux élevé de liaisons réalisées en /n/.

Par ailleurs, tous les contextes identifiés en /n/ sont monosyllabiques. Sachant que ces contextes représentent 85,1% des liaisons réalisées en /n/ (*i.e.* dans *En*) (cf. Figure 6-28 : 220), on peut supposer de façon beaucoup plus générale que ce sont surtout les monosyllabiques qui sont amenés à lier en /n/. Cela suppose également que, si dans *En*, il existe des contextes de liaisons avec des polysyllabiques, l'espace de réalisation de la liaison qu'ils peuvent occuper ne peut excéder les 15 %. Un examen poussé de la base de données PFC montre que la plupart des polysyllabes à finale /n/ qui ont été codés,

sont en fait des noms singuliers, bien connus pour ne pas réaliser la liaison¹¹⁸ (par ex. *en version20 originale* ; *la télévision20 est là* ; *l'inspection20 académique*).

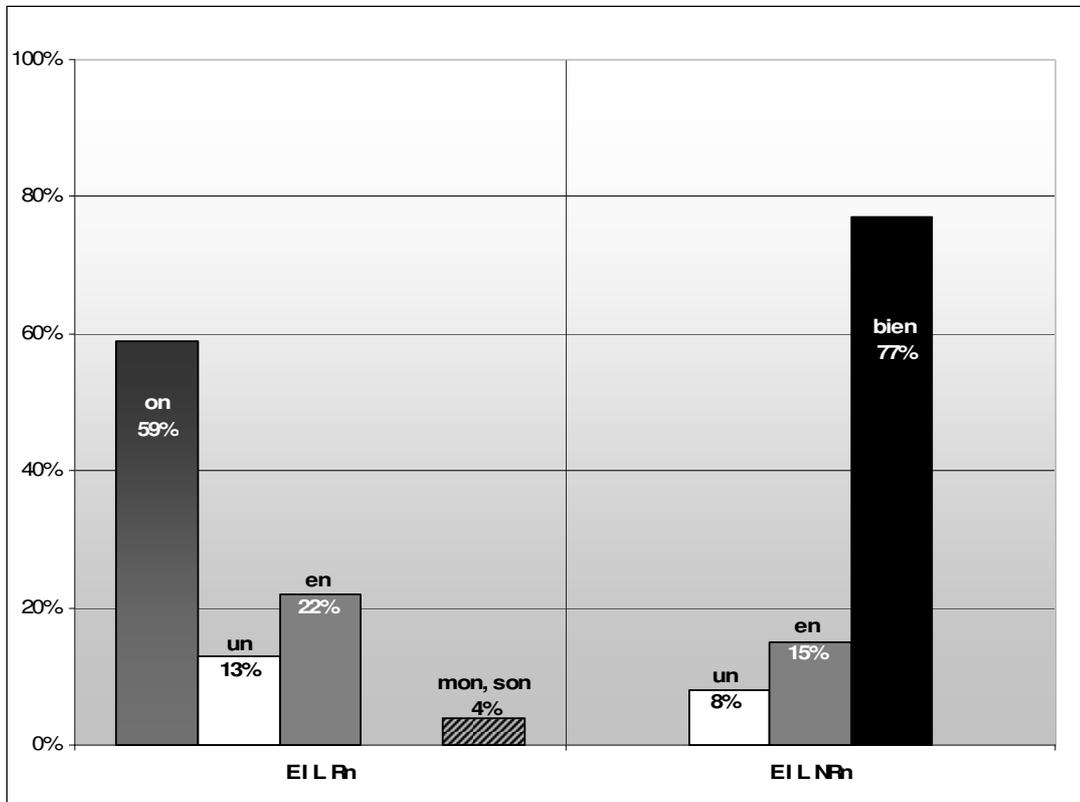


Figure 6-43 Représentation graphique des principales unités constitutives de EILRn et EILNRn

6.3.3.2.2 La consonne /t/

Le nombre de contextes potentiels de liaison avec la consonne /t/ est de 8 113, répartis de la façon suivante :

- 1 188 contextes de liaisons réalisées ;
- et 6 225 contextes de liaisons non réalisées.

¹¹⁸ Si nous n'avons pas tenu compte de ces contextes dans *En*, c'est parce que le protocole PFC ne prévoit pas de codage pour les contextes considérés comme erratiques, exception faite lorsque la CL est réalisée (cf. 5.1.3). Le contexte « Nom sing. + » intègre le paradigme des liaisons erratiques. Aussi, si certains de ces contextes ont fait l'objet d'un codage alors même que la liaison n'avait pas lieu, nous les considérons comme hors de propos puisque relevant d'une « erreur de codage » de la part du transcrip-teur-codeur (cf. 5.1.3.3.1).

La liaison en /t/ est donc réalisée dans 23,27 % des contextes :

	<i>Et</i>	ELRt	ELNRt
Nombre	8 113	1 888	6 225
%	100 %	23,27 %	76,72 %

Figure 6-44 Fréquence des liaisons réalisées et non réalisées dans *Et*

Sur ces 8 113 contextes, nous en avons identifiés 4 972, soit 61,3 % d'entre eux (cf. Figure 6-26) :

- 1 610 contextes de liaisons réalisées ;
- et 3 362 contextes de liaisons non réalisées.

Autrement dit, sur le nombre total de contextes potentiels (*Et*), nous avons identifié 85,3 % des liaisons réalisées et 54 % des liaisons non réalisées. Le tableau ci-dessous synthétise ces résultats :

	<i>Et</i>	<i>EIt</i>	ELRt	EILRt	ELNRt	EILNRt
Nombre	8 113	4 972	1 888	1 610	6 225	3 362
%	100 %	61,28 %	100 %	85,28 %	100 %	54 %

Figure 6-45 Contextes identifiés pour la CL /t/ (*Et*)

De la même façon que pour la CL /n/, les remarques qui vont suivre ne portent que sur les données qui constituent l'ensemble *EIt*.

Dans *EIt*, la liaison est réalisée à hauteur de 32,4 % comme le montre le tableau suivant :

	<i>EIt</i>	EILRt	EILNRt
Nombre	4 972	1 610	3 362
%	100 %	32,38 %	67,61 %

Figure 6-46 Fréquence des liaisons réalisées et non réalisées dans *EIt*

Selon les mots, la liaison est plus ou moins réalisée, comme on peut le voir dans le tableau ci-dessous :

	M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
1	<i>tout</i>	287	238	49	82,9 %
2	<i>quand</i>	639	498	141	77,9 %
3	<i>petit</i>	48	36	12	75 %
4	<i>grand</i>	9	5	4	55,6 %
5	<i>est</i> ¹¹⁹	636	279	357	43,9 %
6	<i>peut</i>	94	29	65	30,9 %
7	<i>c'est</i>	1 470	413	1 057	28,1 %
8	<i>sont</i>	208	40	168	19,2 %
9	<i>étaient</i>	71	8	63	11,3 %
10	<i>ont</i>	178	17	161	9,6 %
11	<i>était</i>	343	29	314	8,5 %
12	<i>allait</i>	45	2	43	4,4 %
13	<i>avait</i>	374	10	364	2,7 %
14	<i>avaient</i>	408	6	402	1,5 %
15	<i>faut</i>	127	0	127	0 %
16	<i>fallait</i>	35	0	35	0 %
	Total	4 972	1 610	3 362	32,4 %

Figure 6-47 M1 contenus dans *Et*

Un inventaire des M1 révèle que d'une part, ces unités sont fort variées et que, d'autre part, elles correspondent le plus souvent à des formes verbales fléchies (par ex. des formes fléchies de *être*, *aller*, *avoir*). Cette diversité fait qu'il nous a été impossible d'identifier plus de contextes contenus dans *Et*, et contribue probablement au fait que la fréquence de réalisation de la liaison soit si variable.

L'observation de ces formes pouvant présenter une liaison en /t/ et leur taux respectif de liaison réalisée indiquent que :

- (i) la fréquence de réalisation de liaison n'est jamais optimale, quelque soit le mot employé ;
- (ii) cette fréquence est plus ou moins importante selon les mots en jeu dans le processus. Les taux les plus faibles s'observent pour les « Verbes + » et les taux de réalisation les plus élevés se retrouvent avec les mots invariables ;

¹¹⁹ Nous n'avons relevé aucune occurrence du nom singulier *est* [est]. Il faut donc entendre le mot *est* comme une forme verbale fléchie du verbe *être*.

(iii) enfin cette fréquence semble indépendante du nombre d'occurrences. Autrement dit, certains mots très représentés, peuvent ne réaliser la liaison que dans une faible proportion, tandis que d'autres, bien moins nombreux en contexte de liaison, ont un taux élevé de réalisation de la liaison.

L'examen que nous opérons dans la suite de cette section s'organise de la façon suivante : (i) les verbes, qui sont analysés selon leur forme de base (cf. 1.) ; (ii) *petit* et *grand* (cf. 2.) ; et (iii.) les mots invariables (cf. 3.).

1. Les verbes

L'analyse des verbes contenant la CL /t/ se fera en trois étapes : (i) les formes fléchies du verbe *être* ; (ii) les formes fléchies du verbe *avoir* ; (iii) les autres formes verbales.

(i) Les formes fléchies de *être*

Sur un total de 2 728 contextes, la liaison ne se réalise que dans 769 contextes, soit 28,2 % (cf. Figure 6-48).

Nous avons identifié 5 constructions avec le verbe *être* :

- le présentatif (*c'est*) :

(79) *c'est11t un métier*

- la 3ème personne du singulier, présent de l'indicatif (*est*) :

(80) *un garçon qui est11t artisan*

- la 3ème personne du pluriel, présent de l'indicatif (*sont*) :

(81) *mes enfants sont11t à Montpellier*

- la 3ème personne du singulier, imparfait (*était*) :

(82) *mon père était21t originaire*

- et la 3ème personne du pluriel, imparfait (*étaient*) :

(83) *ils étaient21t attachés*

La fréquence de réalisation de chacune de ces formes est représentée dans le tableau ci-dessous :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>c'est</i>	1 470	413	1 057	28,1 %
<i>est</i>	636	279	357	43,9 %
<i>sont</i>	208	40	168	19,2 %
<i>était</i>	343	29	314	8,5 %
<i>étaient</i>	71	8	63	11,3 %
Total	2 728	769	1 959	28,19 %

Figure 6-48 **Fréquence des liaisons en /t/ avec les formes fléchies du verbe être**

Les taux de réalisation de la liaison sont différents pour chacune des formes. Le présentatif *c'est* est la forme la plus représentée avec 1 470 occurrences ; la troisième personne du pluriel de l'imparfait est la forme la moins représentée avec 71 occurrences.

La forme « 3ème personne du singulier, présent de l'indicatif » (*est*) est celle qui présente le plus de liaisons réalisées (43,9 %). Le nombre de contextes potentiels de liaison dans lesquels elle apparaît (616 au total) est cependant plus de deux fois moins important que la forme présentative *c'est* qui ne réalise la liaison qu'à hauteur de 28,1 %.

Les taux de réalisation de la liaison avec les trois autres formes sont relativement faibles : 19,2 % pour *sont*, 11,3 % pour *étaient* et 8,5 % pour *était*.

Ensemble, ces formes représentent 54,9 % des contextes de liaison dans *EIt*, 47,8 % des liaisons réalisées dans *EILRt* et 58,3 % des liaisons non réalisées dans *EILNRt* :

	dans <i>Et</i>		dans <i>EILRt</i>		dans <i>EILNRt</i>	
	Nombre	%	Nombre	%	Nombre	%
<i>c'est</i>	1 470	29,57 %	413	25,65 %	1 057	31,44 %
<i>est</i>	636	12,79 %	279	17,33 %	357	10,62 %
<i>sont</i>	208	4,18 %	40	2,48 %	168	5,00 %
<i>était</i>	343	6,90 %	29	1,80 %	314	9,34 %
<i>étaient</i>	71	1,43 %	8	0,50 %	63	1,87 %
Total	2 728 / 4 972	54,87 %	769 / 1 610	47,76 %	1 959 / 3 362	58,27 %

Figure 6-49 Représentation des formes fléchies de *être* dans *Et*, *EILRt* et *EILNRt*

(ii) Les formes fléchies de *avoir*

Sur 960 formes, la liaison n'est réalisée que dans 3,4 % des contextes (cf. Figure 6-50).

Trois formes fléchies du verbe *avoir* ont été identifiées :

- la 3ème personne du pluriel du présent de l'indicatif (*ont*) :

(84) *Les gens ont10 eu l'habitude*

- la 3ème personne du singulier du prétérit (*avait*) :

(85) *il y avait21t aussi*

- et la 3ème personne du pluriel du prétérit (*avaient*) :

(86) *ils avaient20 échangé*

Les fréquences de réalisation de ces trois formes sont exposées dans le tableau ci-dessous :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>ont</i>	178	17	161	9,6 %
<i>avait</i>	374	10	364	2,7 %
<i>avaient</i>	408	6	402	1,5 %
Total	960	33	927	3,44 %

Figure 6-50 Fréquence des liaisons en /t/ avec les formes fléchies du verbe *avoir*

On constate que ces formes sont très peu représentées dans le corpus observé et que, quelle que soit la forme utilisée, la liaison n'est pratiquement jamais réalisée : par exemple, sur 408 occurrences, *avaient* ne réalisent la CL que 6 fois (1,5 %). Au maximum, on observe une fréquence de réalisation de la liaison de 9,6 % pour *ont*, présent dans 178 contextes.

Si l'on regarde maintenant ces formes dans *EIt*, *EILRt* et *EILNRt*, on constate qu'elles représentent :

- 19,3 % des contextes de liaisons en /t/ ;
- 2,05 % des liaisons réalisées ;
- et 27,6 % des non-liaisons en /t/.

	dans <i>EIt</i>		dans <i>EILRt</i>		dans <i>EILNRt</i>	
	Nombre	%	Nombre	%	Nombre	%
<i>ont</i>	178	3,58 %	17	1,06 %	161	4,79 %
<i>avait</i>	374	7,52 %	10	0,62 %	364	10,83 %
<i>avaient</i>	408	8,21 %	6	0,37 %	402	11,96 %
Total	960 / 4 972	19,31 %	33 / 1 610	2,05 %	927 / 3 362	27,57 %

Figure 6-51 Représentation des formes fléchies de *avoir* dans *EIt*, *EILRt* et *EILNRt*

(iii) Les autres formes verbales

Sur 301 contextes, la liaison n'est réalisée que 31 fois, soit 10,3 % (cf. Figure 6-52).

Parmi ces « autres formes verbales » identifiées, on compte :

- le verbe *aller* à la 3ème personne du singulier à l'imparfait (*allait*) :

(87) *ça allait*²⁰ un peu mieux

(88) *il allait*²¹ à son travail

- le verbe *pouvoir* à la 3ème personne du singulier, présent de l'indicatif (*peut*), surtout présent dans le *peut-être* (90) :

(89) *on peut*¹⁰ être

(90) *peut*¹¹ être qu'il faut

- le verbe *falloir* à la 3ème personne du singulier, présent de l'indicatif (*faut*) :

(91) *il faut10 être de gauche*

- et enfin, le verbe *falloir* à la 3ème personne du singulier de l'imparfait (*fallait*) :

(92) *il fallait20 une chambre à deux*

La fréquence de réalisation de la liaison de ces formes est représentée dans le tableau suivant :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>allait</i>	45	2	43	4,4 %
<i>peut</i>	94	29	65	30,9 %
<i>faut</i>	127	0	127	0 %
<i>fallait</i>	35	0	35	0 %
Total	301	31	270	10,30 %

Figure 6-52 Fréquence des liaisons en /t/ avec les autres formes verbales

On note que seul *peut* réalise fréquemment la liaison (30,9 %). La liaison n'est, en revanche, jamais réalisée avec le verbe *falloir*, et très peu avec la forme *allait* (4,4 %).

Comme on peut le constater dans la Figure 6-53, ces formes ne représentent que très peu de contextes dans l'ensemble *EIt* (6 %), que ce soit dans l'ensemble des liaisons non réalisées (8 %) ou plus encore dans l'ensemble des liaisons réalisées (1,9 %).

	dans <i>EIt</i>		dans <i>EILRt</i>		dans <i>EILNRt</i>	
	Nombre	%	Nombre	%	Nombre	%
<i>allait</i>	45	0,91 %	2	0,12 %	43	1,28 %
<i>peut</i>	94	1,89 %	29	1,80 %	65	1,93 %
<i>faut</i>	127	2,55 %	0	0 %	127	3,78 %
<i>fallait</i>	35	0,70 %	0	0 %	35	1,04 %
Total	301 / 4 972	6,05 %	31 / 1 610	1,93 %	270 / 3 362	8,03 %

Figure 6-53 Représentation des autres formes verbales dans *EIt*, *EILRt* et *EILNRt*

2. *Petit et grand*

Les formes *petit* et *grand* sont présentes respectivement dans 48 et 9 contextes de liaison. La liaison est fréquente, comme on peut le voir dans le tableau ci-dessous :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>petit</i>	48	36	12	75 %
<i>grand</i>	9	5	4	55,6 %

Figure 6-54 Fréquence des liaisons en /t/ avec *petit* et *grand*

Si l'on observe leur représentation dans les ensembles *EIt*, *EILRt* et *EILNRt*, on constate que :

- *petit* représente :
 - 0,96 % des contextes de *EIt* ;
 - 2,24 % des contextes de liaisons réalisées en *EILRt* :

(93) *son petit21 enfant*

- et 0,36 % des contextes de liaisons non réalisées *EILNRt* :

(94) *un endroit assez petit20 en fait*

Ces résultats sont présentés dans le tableau suivant :

<i>petit</i>	dans <i>EIt</i>	dans <i>EILRt</i>	dans <i>EILNRt</i>
Nombre	48 / 4 972	36 / 1 610	12 / 3 362
%	0,96 %	2,24 %	0,36 %

Figure 6-55 Représentation de *petit* dans *EIt*, *EILRt* et *EILNRt*

- et *grand* représente quant à lui :
 - 0,18 % des contextes de *EIt* ;
 - 0,31 % des contextes de *EILRt* :

(95) *Le grand11 Ouest*

- et 0,12 % des contextes de *EILNRt* :

(96) *le grand10 aussi*

Le tableau ci-dessous représente ces résultats :

<i>grand</i>	dans <i>EIt</i>	dans <i>EILRt</i>	dans <i>EILNRt</i>
Nombre	9 / 4 972	5 / 1 610	4 / 3 362
%	0,18 %	0,31 %	0,12 %

Figure 6-56 Représentation de *grand* dans *EIt*, *EILRt* et *EILNRt*

Le faible taux de réalisation de la liaison avec *tout* s'explique en partie par sa présence dans la forme de négation « *pas du tout* », dans laquelle *tout* ne réalise jamais la liaison (par ex. *pas du tout10 apprécié* ; *pas du tout10 un public*).

3. *Tout*

Tout est présent dans 287 contextes :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>tout</i>	287	238	49	82,9 %

Figure 6-57 Fréquence des liaisons en /t/ avec *tout*

Il est identifié dans :

- 5,8 % des contextes de liaisons en /t/ (soit *EIt*) ;
- 14,8 % des contextes de liaisons réalisées en /t/ (soit *EILRt*) :

(97) *j'ai tout11t arraché*

- et 1,5 % des contextes de liaisons non réalisées en /t/ (soit *EILNRt*) :

(98) *presque tout10 à la main*

(99) *j'ai pas du tout10 apprécié*

Ces résultats sont représentés dans le tableau ci-dessous :

<i>tout</i>	dans <i>EIt</i>	dans <i>EILRt</i>	dans <i>EILNRt</i>
Nombre	287 / 4 972	238 / 1 610	49 / 3 362
%	5,77 %	14,78 %	1,46 %

Figure 6-58 Représentation de *tout* dans *EIt*, *EILRt* et *EILNRt*

Le faible taux de réalisation de la liaison avec *tout* s'explique en partie par sa présence dans l'expression « *pas du tout* », laquelle ne constitue pas un contexte dans lequel la liaison est réalisée (par ex. *pas du tout10 apprécié* ; *pas du tout10 un public*).

4. *Quand*

Quand est présent dans 639 contextes de liaison ; la liaison n'est pas toujours réalisée :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>quand</i>	639	498	141	77,9 %

Figure 6-59 Fréquence des liaisons en /t/ avec *quand*

Quand est présent dans :

- 12,9 % des contextes de *EIt* ;
- 30,9 % des contextes de *EILRt* :

(100) *quand11t il est rentré*

- et 4,2 % des contextes de *EILNRt* :

(101) *quand10 on fait une adaptation*

Soit :

<i>quand</i>	dans <i>EIt</i>	dans <i>EILRt</i>	dans <i>EILNRt</i>
Nombre	639 / 4 972	498 / 1 610	141 / 3 362
%	12,85 %	30,93 %	4,19 %

Figure 6-60 Représentation de *quand* dans *EIt*, *EILRt* et *EILNRt*

5. Bilan

Dans le tableau ci-dessous, nous reprenons les fréquences de réalisation de la liaison dans *EIn*. Nous avons regroupé dans un même ensemble les formes de *être* (être), de *avoir* (avoir), et les autres formes verbales (autres verbes) :

<i>/t/</i>	EIt	EILRt	EILNRt
être	54,9 %	47,8 %	58,3 %
avoir	19,3 %	2,1 %	27,6 %
autres verbes	6,1 %	1,9 %	8 %
<i>petit</i>	1 %	2,2 %	0,4 %
<i>grand</i>	0,2 %	0,3 %	0,1 %
<i>tout</i>	5,8 %	14,8 %	1,5 %
<i>quand</i>	12,9 %	30,9 %	4,2 %
Total	100 %	100 %	100 %

Figure 6-61 Fréquence de réalisation de la liaison dans *EIt*

Ce tableau indique que :

- *EILRt* est principalement constitué des formes fléchies de *être* (47,8 %), des contextes « *quand +* » (30,9 %) et des contextes « *tout* » (14,8 %) ;
- *EILNRt* est principalement constitué des formes fléchies de *être* (58,3 %), des formes fléchies de *avoir* (27,6 %), des autres formes verbales (8 %) et de s contextes « *quand +* » (4,2 %).

Nous proposons ci-dessous une représentation graphique des principales formes constitutives des ensembles *EILRt* et *EILNRt* :

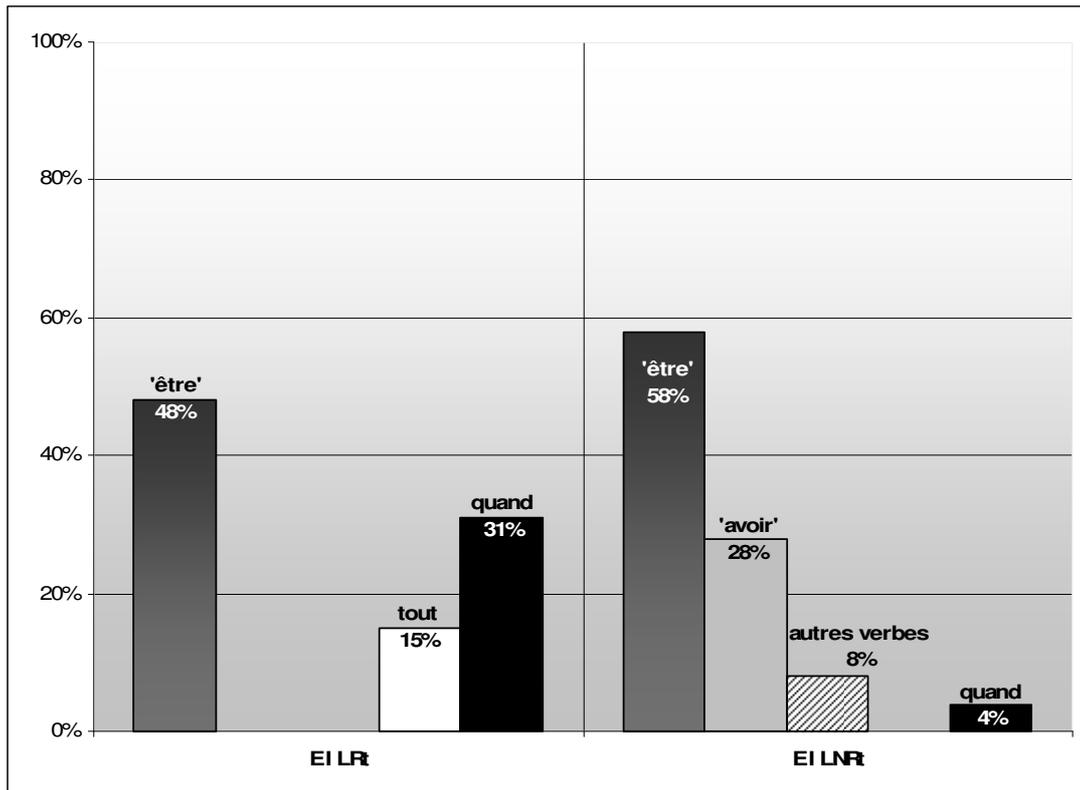


Figure 6-62 Représentation graphique des principales unités constituantes de EILRt et EILNRt

On voit que les différents contextes contenant une forme du verbe *être* sont très représentés dans *EIt* (2 728 au total) et manifestent une liaison variable. Cette liaison est plus ou moins variable selon la forme et est comprise entre 0,5 % et 25,6 % de réalisation (cf. Figure 6-49). Les taux les plus faibles s'observent pour les formes dont le nombre d'occurrences est le moins élevé (*étaient* : 71 occurrences, 8 réalisations), le taux le plus fort, pour *c'est* dont le nombre d'occurrences est le plus élevé (1 470 au total).

Les contextes de liaison avec une forme fléchie du verbe *avoir* sont beaucoup moins nombreux (960 au total) et ne liaisonnent pratiquement pas : les taux de réalisation varient entre 0,4 % et 1,1 % (cf. Figure 6-51). Comparativement aux formes fléchies du verbe *être*, le nombre d'occurrences est faible (960 contre 2 728).

Concernant les « autres » formes verbales, qui représentent 8 % de EILNRt, on constate que seul *peut* et *allait* « réalisent » la liaison (respectivement 1,8 % et 0,1 % ;

cf. Figure 6-53). Une analyse des contextes de *peut* montre qu'en fait, la liaison n'apparaît qu'à l'intérieur du groupe figé « *peut-être* », les autres contextes ne la réalisant jamais.

Concernant maintenant *petit* et *grand* : la liaison est variable. Une analyse plus précise des contextes dans lesquels *petit* et *grand* apparaissent, indique que lorsque *petit* est employé comme simple adjectif qualificatif antéposé (par ex. *petit ami*), la liaison est catégorique. Par contre, lorsqu'il est postposé (par ex. *un endroit assez petit²⁰ en fait*) ou qu'il se trouve en deuxième position du groupe figé *petit à petit* (par ex. *petit à petit²⁰ elle lit*), alors la liaison est erratique. De la même façon, lorsque *grand* est utilisé comme simple adjectif qualificatif, la liaison est catégorique (par ex. *un grand¹¹ oral*) ; par contre, il ne liera pas, lorsqu'il sera employé comme nom (par ex. *le grand¹⁰ aussi*).

On peut être surpris par le très faible nombre d'occurrences de *petit* et *grand* en contexte de liaison, étant donné qu'ils sont décrits par la littérature comme étant très fréquents. Cette remarque nécessite toutefois la prudence étant donné que nous n'observons ces termes que dans un contexte de liaison.

Aussi, si l'on identifie ces mots en les regroupant uniquement en fonction de leur CL, sans vérifier leurs différents emplois, on en conclut que la liaison qu'ils manifestent est variable. Si l'on procède à une analyse plus fine, on s'aperçoit que ces mots manifestent tel ou tel type de liaison en fonction de leur emploi, bien qu'ils conservent toujours la même consonne finale.

6.3.3.2.3 La consonne /z/

Le nombre potentiel de contextes de liaison possibles avec la consonne /z/ est de 13 852, répartis de la façon suivante :

- 5 960 contextes de liaisons réalisées ;
- et 7 892 contextes de liaisons non réalisées.

La liaison en /z/ est donc effective dans 43 % des contextes :

	<i>Ez</i>	ELRz	ELNRz
Nombre	13 852	5 960	7 892
%	100 %	43,02 %	56,97 %

Figure 6-63 Fréquence des liaisons réalisées et non réalisées dans *Ez*

La très forte représentation de la CL /z/ est principalement liée au fait que cette consonne qui apparaît orthographiquement sous la forme {s, z ou x} est souvent placée à la finale des mots pluriels (déterminants, noms) et aussi de quelques mots invariables (cf. Figure 6-66). Comparativement aux CL /n/ et /t/, les mots à finale /z/ sont plus nombreux dans le corpus : 28 au total ont été recensés dans *EIz* (cf. Figure 6-66).

La très forte variation de la fréquence d'apparition de /z/ est liée au fait que les occurrences dans lesquelles la CL apparaît correspondent le plus souvent à des contextes de liaisons variables. Etant donné la multiplicité de ces formes réalisant une liaison variable, nous en avons identifié un peu de moins de 50 %.

Aussi, sur l'ensemble des 13 852 contextes de *Ez*, nous avons identifié 8 437 contextes, soient 60,9 %, répartis de la façon suivante (cf. Figure 6-26) :

- 4 984 contextes de liaisons réalisées ;
- et 3 453 contextes de liaisons non réalisées.

A l'intérieur de *Ez*, nous disposons donc de 83,6 % des contextes de liaisons réalisées et 43,8 % des non-liaisons qui auraient pu être réalisées en /z/. Le tableau ci-dessous synthétise ces résultats :

	<i>Ez</i>	<i>EIz</i>	ELRz	EILRz	ELNRz	EILNRz
Nombre	13 852	8 437	5 960	4 984	7 892	3 453
%	100 %	60,90 %	100 %	83,62 %	100 %	43,75 %

Figure 6-64 Contextes identifiés pour la CL /z/ (*Ez*)

Dans cet ensemble (*EIz*), la liaison est réalisée dans 59,1 % des contextes :

	<i>EIz</i>	<i>EILRz</i>	<i>EILNRz</i>
Nombre	8 437	4 984	3 453
%	100 %	59,07 %	40,92 %

Figure 6-65 **Fréquence des liaisons réalisées et non réalisées dans *EIz***

L'examen de ces formes (*cf.* Figure 6-66) montre que :

- certaines réalisent toujours la liaison (par ex. les déterminants *les* et *ses*) ;
- d'autres jamais (par ex. *toujours* et *ans*)
- pour certains mots, la liaison est plus ou moins réalisée (par ex. la liaison est réalisée dans près de 75 % des contextes avec *nous*, alors qu'elle n'est réalisée que dans 13,5 % des contextes avec *suis*) ;
- et cette fréquence, plus ou moins importante, semble indépendante du nombre de contextes dans lesquels le mot apparaît (par ex. *toujours* apparaît dans 184 contextes de liaison et ne réalise jamais la liaison, alors que *très*, présent dans 145 contextes, réalise la CL dans 96,6 % des cas).

	M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
1	<i>deux</i>	334	334	0	100 %
2	<i>mes</i>	148	148	0	100 %
3	<i>ses</i>	62	62	0	100 %
4	<i>ils</i>	864	853	11	98,73 %
5	<i>les</i>	1 047	1 030	17	98,38 %
6	<i>vous</i>	433	425	8	98,15 %
7	<i>des</i>	796	769	27	96,61 %
8	<i>très</i>	145	140	5	96,55 %
9	<i>dans</i>	398	378	20	94,97 %
10	<i>petites</i>	13	12	1	92,31 %
11	<i>trois</i>	211	182	29	86,26 %
12	<i>chez</i>	64	53	11	82,81 %
13	<i>grandes</i>	5	4	1	80 %
14	<i>nous</i>	427	320	107	74,94 %
15	<i>grands</i>	7	5	2	71,43 %
16	<i>plus</i>	248	159	89	64,11 %
17	<i>sommes</i> ¹²⁰	27	12	15	44,44 %
18	<i>petits</i>	22	9	13	40,91 %
19	<i>êtes</i>	16	5	11	31,25 %
20	<i>avons</i>	24	5	19	20,83 %
21	<i>suis</i>	430	58	372	13,49 %
22	<i>assez</i>	41	2	39	4,88 %
23	<i>vais</i>	51	2	49	3,92 %
24	<i>pas</i> ¹²¹	880	12	868	1,36 %
25	<i>mais</i>	791	4	787	0,51 %
26	<i>puis</i>	639	1	638	0,16 %
27	<i>toujours</i>	184	0	184	0 %
28	<i>ans</i>	130	0	130	0 %
	Total	8 437	4 984	3 453	59,07 %

Figure 6-66 M1 contenus dans *Elz*

¹²⁰ Dans le corpus étudié, *sommes* correspond toujours à la forme verbale fléchie de *être* et jamais à la forme plurielle du nom *somme*.

¹²¹ De la même façon, le mot *pas* ne correspond jamais au nom singulier ou pluriel mais renvoie toujours à un adverbe.

Pour examiner ces mots, nous constituons plusieurs groupes :

1. les déterminants *mes, ses, des* et *les* ;
2. les numéraux *deux* et *trois* ;
3. le groupe des pronoms
4. le groupe des formes verbales *être, avoir* et *aller* ;
5. le groupe des « invariables » ;
6. et le groupe des « autres » formes contenant la CL /z/

1. Les déterminants

Tous les déterminants contenant la CL /z/ sont au pluriel comme en atteste le tableau ci-dessous :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>les</i>	1 047	1 030	17	98,38 %
<i>des</i>	796	769	27	96,61 %
<i>mes</i>	148	148	0	100 %
<i>ses</i>	62	62	0	100 %
Total	2 053	2 009	44	97,31 %

Figure 6-67 Fréquence des liaisons en /z/ avec les déterminants

Ces déterminants sont au nombre de quatre :

1. l'article pluriel *les* :

(102) *il fait les11z ouvertures*

(103) *c'était les10 hiboux*

(104) *les10 un an*

2. l'article pluriel *des* :

(105) *il y a des11z accents qui me dérangent*

(106) *des10 hauts et des bas*

(107) *des10 hurlements*

3. le possessif *mes* :

(108) *j'ai fini mes11z exams le huit*

4. le possessif *ses* :

(109) *une fille qui [...] finit ses11z humanités*

Seuls deux de ces déterminants ne réalisent pas la liaison à 100 % : *les* (102) et *des* ((105) et (107)). La liaison n'est en effet pas réalisée dans le contexte « + h aspiré », qui est défini comme erratique (cf. Figure 3-12, p. 78). On compte également un contexte dans lequel *les* précède l'expression « *un an* » (104).

Les déterminants identifiés représentent 24,3 % des contextes de liaisons dans EIZ, 40,3 % des liaisons réalisées dans EILRz et seulement 1,3 % des liaisons non réalisées dans EILNRz :

	dans EIZ		dans EILRz		Dans EILNRz	
	Nombre	%	Nombre	%	Nombre	%
<i>mes</i>	148	1,75 %	148	2,97 %	0	0 %
<i>mes</i>	62	0,73 %	62	1,24 %	0	0 %
<i>des</i>	796	9,43 %	769	15,43 %	27	0,78 %
<i>les</i>	1 047	12,41 %	1 030	20,67 %	17	0,49 %
Total	2 053 / 8 437	24,33 %	2 009 / 4 984	40,31 %	44 / 3 453	1,27 %

Figure 6-68 Représentation des déterminants dans EIZ, EILRz et EILNRz

2. *Deux et trois*

On retrouve les formes *deux* et *trois* dans respectivement 334 et 211 contextes. La liaison avec *deux* est toujours réalisée, mais seulement fréquente avec *trois* :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>deux</i>	334	334	0	100 %
<i>trois</i>	211	182	29	86,26 %
Total	545	516	29	94,68 %

Figure 6-69 Fréquence des liaisons en /z/ avec *deux* et *trois*

La Figure 6-70 montre que si *deux* et *trois* ne représentent que peu de contextes dans l'ensemble *EIz*, ils représentent toutefois 10,4 % de l'ensemble des liaisons réalisées dans EILRz :

1. *deux* est présent dans :

- 3,96 % des contextes de *EIz* ;
- 6,70 % des contextes de EILRz :

(110) *on avait deux11z heures de néerlandais*

- mais dans aucun contexte de liaison non réalisée (EILNRz).

2. *trois* est présent dans :

- 2,5 % des contextes de *EIz* ;
- 3,7 % des contextes de EILRz ;
- et 0,84 % des contextes de EILNRz :

(111) *ces trois10 ou quatre documents*

(112) *qui se mettent deux trois10 ensemble*

	dans <i>EIz</i>		dans EILRz		Dans EILNRz	
	Nombre	%	Nombre	%	Nombre	%
<i>deux</i>	334	3,96 %	334	6,70 %	0	0 %
<i>trois</i>	211	2,50 %	182	3,65 %	29	0,84 %
Total	545 / 8 437	6,46 %	516 / 4 984	10,35 %	29 / 3 453	0,84 %

Figure 6-70 Représentation de *deux* et *trois* dans *EIz*, EILRz et EILNRz

3. *Les pronoms*

La liaison avec les pronoms est très fréquemment réalisée. Les pronoms identifiés sont au nombre de trois :

1. la 3^{ème} personne du pluriel (*ils*) :

(113) *ils11z étaient à moitié enterrés*

(114) *ils10 en ressortent tout habillé*

2. la 2^{ème} personne du pluriel (*vous*) :

(115) *quand vous11z allez toucher la main*

(116) *vous10 aussi*

3. et la 1^{ère} personne du pluriel (*nous*) :

(117) *on nous11z avait donné les primes d'arrachage*

(118) *parce que nous10 on vote toujours*

Les fréquences de réalisation de ces pronoms sont présentées dans le tableau suivant :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>ils</i>	864	853	11	98,73 %
<i>vous</i>	433	425	8	98,15 %
<i>nous</i>	427	320	107	74,94 %
Total	1 724	1 598	126	92,69 %

Figure 6-71 **Fréquence des liaisons en /z/ avec les pronoms**

Ils, *nous* et *vous*, pronoms sujets, manifestent toujours une liaison. Toutefois, il apparaît important de préciser que la liaison n'est pas toujours réalisée lorsque *nous* et *vous* sont (i) détachés (118) et (ii) suivis par « *aussi* » (116).

L'observation de ces pronoms dans EIZ, EILRz et EILNRz indique que ces formes représentent :

- 20,4 % de l'ensemble des contextes de EIZ ;
- 32,1 % des liaisons (de EILRz) ;
- et 3,7 % des non-liaisons (EILNRz) :

	dans EIZ		Dans EILRz		dans EILNRz	
	Nombre	%	Nombre	%	Nombre	%
<i>ils</i>	864	10,24 %	853	17,11 %	11	0,32 %
<i>vous</i>	433	5,13 %	425	8,53 %	8	0,23 %
<i>nous</i>	427	5,06 %	320	6,42 %	107	3,10 %
Total	1 724 / 8 437	20,43 %	1 598 / 4 984	32,06 %	126 / 3 453	3,65 %

Figure 6-72 Représentation des pronoms dans EIZ, EILRz et EILNRz

4. Les verbes

L'analyse des formes verbales contenant la CL /z/ se fera en deux temps : (i) les formes fléchies du verbe *être* ; et les autres formes verbales.

(i) Les formes fléchies de *être*

La liaison avec les différentes formes du verbe *être* est peu réalisée : 15,9 % de réalisation (cf. Figure 6-73).

Trois formes fléchies ont été identifiées :

1. 1ère personne du pluriel, présent de l'indicatif (*sommes*) :

(119) *nous sommes*₂₀ arrivés à Douzens

(120) *nous sommes*_{21z} allés à Montpellier

2. 2ème personne du pluriel, présent de l'indicatif (*êtes*) :

(121) *vous êtes*₁₀ en situation difficile

(122) *vous êtes*_{11z} obligé de faire ça

3. 1ère personne du singulier, présent de l'indicatif (*suis*) :

(123) *je suis*_{11z} arrivé dans la salle d'opération

(124) *je suis*¹⁰ *allée donc euh*

La fréquence de réalisation de chacune de ces formes est représentée ci-dessous :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>sommes</i>	27	12	15	44,44 %
<i>êtes</i>	16	5	11	31,25 %
<i>suis</i>	430	58	372	13,49 %
Total	473	75	398	15,86 %

Figure 6-73 Fréquence des liaisons en /z/ avec les formes fléchies de *être*

Les taux de réalisation de la liaison s'échelonnent entre 13,5 % (*suis*) et près de 45 % (*sommes*) pour un nombre total d'occurrences très différents : *suis* est plus de 25 fois plus présent que *êtes* (430 occurrences contre 16) et son taux de réalisation de la CL deux fois moins important.

Ces formes représentent seulement 5,6 % de l'ensemble des contextes de *EIz* et 11,5 % des liaisons non réalisées dans *EILNRz* :

	dans <i>EIz</i>		dans <i>EILRz</i>		dans <i>EILNRz</i>	
	Nombre	%	Nombre	%	Nombre	%
<i>sommes</i>	27	0,32 %	12	0,24 %	15	0,43 %
<i>êtes</i>	16	0,19 %	5	0,10 %	11	0,32 %
<i>suis</i>	430	5,10 %	58	1,16 %	372	10,77 %
Total	473 / 8 437	5,61 %	75 / 4 984	1,50 %	398 / 3 453	11,52 %

Figure 6-74 Représentation des formes fléchies de *être* dans *EIz*, *EILRz* et *EILNRz*

(ii) Les autres formes verbales

Deux autres formes verbales ont été relevées : *avons* (forme fléchie de *avoir*) et *vais* (forme fléchie de *aller*) :

(125) *nous avons*^{21z} *eu une politique agricole*

(126) *nous avons*²⁰ *acheté la petite maison*

(127) *quand je vais11z au cimetière*

(128) *et je vais10 essayer de les retrouver*

Les résultats indiquent que ces formes sont (i) présentes dans très peu de contextes et (ii) liées très peu :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>avons</i>	24	5	19	20,83 %
<i>vais</i>	51	2	49	3,92 %
Total	75	7	68	9,33 %

Figure 6-75 Fréquence des liaisons en /z/ avec les autres formes verbales

Elles ne sont que très peu représentatives des ensembles *EIz* (0,9 %), *EILRz* (0,15 %) et *EILNRz* (2 %), comme on peut le voir dans le tableau suivant :

	dans <i>EIz</i>		Dans <i>EILRz</i>		dans <i>EILNRz</i>	
	Nombre	%	Nombre	%	Nombre	%
<i>avons</i>	24	0,28 %	5	0,10 %	19	0,55 %
<i>vais</i>	51	0,60 %	2	0,04 %	49	0,42 %
Total	473 / 8 437	0,89 %	7 / 4 984	0,14 %	398 / 3 453	1,97 %

Figure 6-76 Représentation des autres formes verbales dans *EIz*, *EILRz* et *EILNRz*

5. Les mots invariables

La liaison avec les « invariables » est peu élevée : 22,1 % (cf. Figure 6-77), alors que le nombre de formes (9 au total) et le nombre de contextes relevés (3 390) le sont.

Tous ces « invariables » relèvent de classe fermées (conjonctions, prépositions, adverbes monosyllabiques). Les taux de réalisation de la CL sont très différents selon les formes et varient entre 0 % et 96,6 % :

- 4 formes ont des fréquences de réalisation de la CL élevées : *très* (96,6 %), *dans* (95 %), *chez* (82,8 %) et *plus* (64,1 %) ;

- pour les cinq autres formes (*assez, pas, mais, puis* et *toujours*), les fréquences de réalisation de la CL n'excède pas 4,9 %.

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>très</i>	145	140	5	96,55 %
<i>dans</i>	398	378	20	94,97 %
<i>chez</i>	64	53	11	82,81 %
<i>plus</i>	248	159	89	64,11 %
<i>assez</i>	41	2	39	4,88 %
<i>pas</i>	880	12	868	1,36 %
<i>mais</i>	791	4	787	0,51 %
<i>puis</i>	639	1	638	0,16 %
<i>toujours</i>	184	0	184	0 %
Total	3 390	749	2 641	22,09 %

Figure 6-77 **Fréquence des liaisons en /z/ avec les mots invariables**

Ces résultats semblent sans rapport (i) avec le nombre de contextes, puisque par exemple, *très*, avec seulement 145 contextes, réalise la liaison dans 96,6 % des cas, alors que *puis*, avec 639 contextes, ne la réalise que dans 0,16 % des cas ; (ii) avec la longueur des formes : exception faite de *toujours*, toutes les formes observées sont monosyllabiques.

Globalement, ces formes représentent 40,2 % de l'ensemble des contextes compris dans *Elz* et seulement 15 % des liaisons réalisées dans *EILRz*. Par contre, elles représentent 76,5 % des liaisons non réalisées en /z/.

Le tableau ci-dessous présente l'ensemble de ces résultats :

	dans <i>EIz</i>		dans <i>EILRz</i>		dans <i>EILNRz</i>	
	Nombre	%	Nombre	%	Nombre	%
<i>très</i>	145	1,72 %	140	2,81 %	5	0,14 %
<i>dans</i>	398	4,72 %	378	7,58 %	20	0,58 %
<i>chez</i>	64	0,76 %	53	1,06 %	11	0,32 %
<i>plus</i>	248	2,94 %	159	3,19 %	89	2,58 %
<i>assez</i>	41	0,49 %	2	0,04 %	39	1,13 %
<i>pas</i>	880	10,43 %	12	0,24 %	868	25,14 %
<i>mais</i>	791	9,38 %	4	0,08 %	787	22,79 %
<i>puis</i>	639	7,57 %	1	0,02 %	638	18,48 %
<i>toujours</i>	184	2,18 %	0	0,00 %	184	5,33 %
Total	3 390 / 8 437	40,18 %	749 / 4 984	15,03 %	2 641 / 3 453	76,48 %

Figure 6-78 Représentation des mots invariables dans *EIz*, *EILRz* et *EILNRz*

On remarque que ce sont surtout *pas*, *mais* et *puis* qui représentent le plus de contextes dans *EILNRz* et qui, *de facto*, contribuent au fait que les « invariables » constituent une part aussi importante de *EILNRz*.

Dans l'ensemble *EILRz*, seul *dans* est vraiment représentatif puisqu'il représente 7,6 % de ces contextes.

6. Les autres mots à CL /z/

Les autres mots à CL /z/ identifiés sont :

1. la forme plurielle du nom *an* (*ans*), pour laquelle la liaison n'est jamais réalisée :

(129) *deux ans*¹⁰ à Nancy

(130) *dix-sept ans*¹⁰ en trente-neuf

2. deux formes de *petit* (*petits* et *petites*) pour lesquelles la liaison est fréquemment réalisée :

(131) *petits*²⁰ on s'entendait bien

(132) *petits*^{21z}-enfants

(133) *quand on était petites*²⁰ on

(134) *des petites21z anecdotes*

3. et deux formes de *grand* (*grands* et *grandes*), pour lesquelles la liaison est également fréquemment réalisée :

(135) *grandes21z écoles*

(136) *les plus grands20 on*

La fréquence de réalisation de ces cinq formes est présentée dans le tableau suivant :

M1	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
<i>ans</i>	130	0	130	0 %
<i>petits</i>	22	9	13	40,91 %
<i>petites</i>	13	12	1	92,31 %
<i>grands</i>	7	5	2	71,43 %
<i>grandes</i>	5	4	1	80 %
Total	177	30	147	1,69 %

Figure 6-79 **Fréquence des liaisons en /z/ avec les autres mots**

Petit(e)s et *grand(e)s* sont présents dans très peu de contextes : *petits* est le mot le plus représenté avec 22 contextes et 40,9 % de LR ; *grandes* est le moins représenté avec 5 contextes et une seule liaison.

Comme on peut le voir dans la Figure 6-80, ces mots ne représentent donc que très peu des formes identifiées à l'intérieur de *EIz*, *EILRz* et *EILNRz* :

- 2,1% de l'ensemble des contextes identifiés ;
- 0,6 % des liaisons réalisées ;
- et 4,3 % des liaisons non réalisées.

	dans <i>EIz</i>		Dans <i>EILRz</i>		dans <i>EILNRz</i>	
	Nombre	%	Nombre	%	Nombre	%
<i>ans</i>	130	1,54 %	0	0 %	130	3,76 %
<i>petits</i>	22	0,26 %	9	0,18 %	13	0,38 %
<i>petites</i>	13	0,15 %	12	0,24 %	1	0,03 %
<i>grands</i>	7	0,08 %	5	0,10 %	2	0,06 %
<i>grandes</i>	5	0,06 %	4	0,08 %	1	0,03 %
Total	177 / 8 437	2,09 %	30 / 4 984	0,6 %	147 / 3 453	4,26 %

Figure 6-80 Représentation des autres mots dans *EIz*, *EILRz* et *EILNRz*

7. Bilan

L'analyse de la réalisation de la liaison avec la CL /z/ fait apparaître que (i) le nombre de formes impliquées est très élevé ; que (ii) les taux de réalisation sont très variés selon la forme.

Dans le tableau ci-dessous, nous reprenons les fréquences de réalisation de la liaison dans *EIz*. Nous avons regroupé les résultats obtenus pour les formes fléchies de *être*, pour les déterminants, les pronoms, les autres verbes, les invariables et les « autres mots ».

<i>/z/</i>	EIz	EILRz	EILNRz
déterminants	24,3 %	40,3 %	1,3 %
<i>deux</i> et <i>trois</i>	6,5 %	10,4 %	0,8 %
pronoms	20,4 %	32,1 %	3,7 %
<i>être</i>	5,6 %	1,5 %	11,5 %
autres verbes	0,9 %	0,1 %	2 %
invariables	40,2 %	15 %	76,5 %
autres mots	2,1 %	0,6 %	4,3 %
Total	100 %	100 %	100 %

Figure 6-81 Fréquence de réalisation de la liaison des mots dans *EIz*

Ce tableau montre que :

- les déterminants et les pronoms représentent la plupart des contextes dans lesquels la liaison est réalisée (respectivement 40,3 % et 32,1 %) ;

- les invariables représentent la plupart des contextes non réalisés (76,2 %). Mais comme nous l'avons vu, ceci est particulièrement vrai dans les cas de *pas*, *mais* et *puis*.

Nous proposons ci-dessous une représentation graphique des principales formes constitutives des ensembles EILRz et EILNRz :

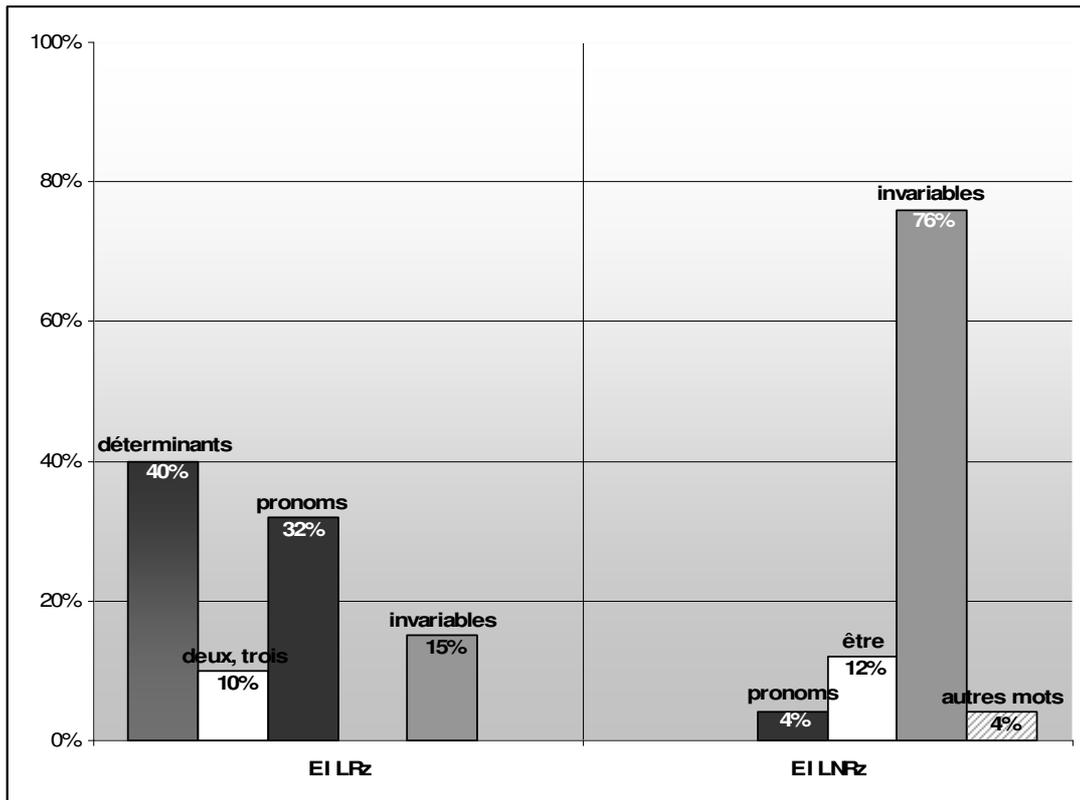


Figure 6-82 Représentation graphique des principales unités constitutives de EILRz et EILNRz

6.3.3.3 Bilan

Dans cette section, nous avons établi qu'effectivement les CL /p, k, r/ sont très peu représentées en contexte de liaison contrairement aux CL /n, t, z/. De ce fait, nous les avons intégrées dans les « consonnes de liaisons non considérées », c'est-à-dire les contextes qui ne seront pas exploités dans les analyses ultérieures.

Concernant maintenant l'ordre de distribution de /n/, /t/ et /z/, nos analyses permettent de déterminer que /n/ est la CL qui réalise le plus la liaison :

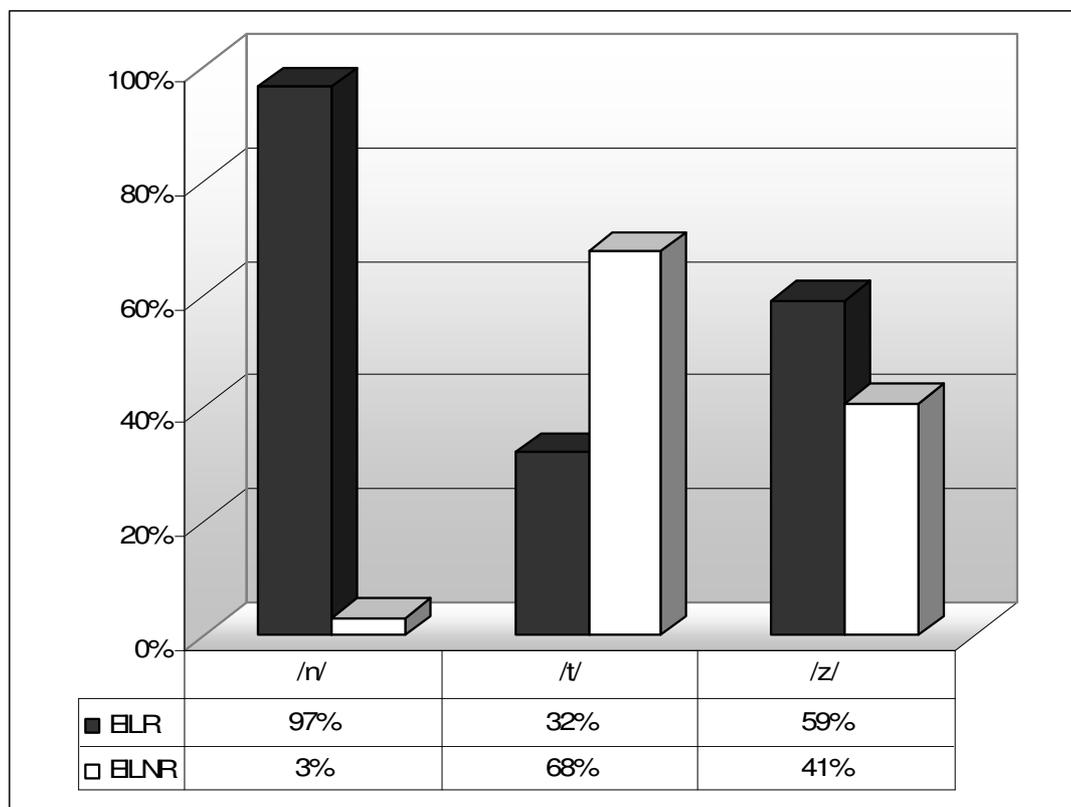


Figure 6-83 Représentation graphique de la réalisation de la liaison des CL /n/, /t/ et /z/ dans EI

L'examen des formes indique que :

- la liaison en /n/ est réalisée à 96,6 % avec 6 formes ;
- la liaison en /z/ est réalisée à 59,1 % avec 28 formes ;
- et que la liaison en /t/ est réalisée à 32,4 % avec 16 formes.

On obtient donc la distribution des CL suivante :

/n/	>	/z/	>	/t/
96,6 % LR		59,1 % LR		32,4 % LR

Figure 6-84 Distribution des consonnes /n/, /z/ et /t/

Si l'on considère maintenant les mots en jeu dans le processus, on rencontre un certain nombre de problèmes relevant de la phonologie (autres que CL) et du lexique :

1. La fréquence de réalisation de la liaison des CL semble influencée par la surreprésentation de certains mots

Nous avons vu que les ensembles *EIn* et *EILRn* sont surreprésentés par *on* qui occupe non seulement 57 % de l'ensemble *EIn*, mais aussi 58,9 % de l'ensemble *EILRn*, puisque sa CL est réalisée dans tous les contextes. Aussi, comparativement aux possessifs *mon* et *son*, dont le nombre de contextes est beaucoup plus restreint (respectivement environ 3 % et 1,3 % des contextes de *EIn* et des contextes de *EILRn*), cette surreprésentation favorise probablement un taux élevé de liaisons réalisées en /n/.

2. Le problème de la longueur

L'analyse de /n/ permet également de poser la question de l'influence de la longueur du mot liaisonnant sur la réalisation de la liaison. On remarque en effet que tous les contextes identifiés en /n/ sont monosyllabiques et que la liaison est réalisée quasi-systématiquement, contrairement aux formes identifiées pour /t/ et /z/ qui peuvent être monosyllabiques ou polysyllabiques et qui ne réalisent pas systématiquement la liaison.

Sachant que ces contextes, monosyllabiques, représentent 85,1 % des liaisons réalisées en /n/ (*i.e.* dans *En* ; cf. Figure 6-28 : 220), on peut supposer de façon beaucoup plus générale que ce sont surtout les monosyllabiques qui sont amenés à lier en /n/. Cela suppose également que, si dans *En*, il existe des contextes de liaisons avec des polysyllabiques, l'espace de réalisation de la liaison qu'ils peuvent occuper ne peut excéder les 15 %. Un examen poussé de la base de données PFC montre que la plupart des polysyllabes à finale /n/ qui ont été codés, sont en fait des noms singuliers, bien connus pour ne pas réaliser la liaison (par ex. *en version20 originale* ; *la télévision20 est là* ; *l'inspection20 académique*).

Par ailleurs, l'analyse de /z/ pose le problème de la longueur ici de manière phonologique et non plus par rapport à la forme écrite relevée. En effet, un certain nombre de mots contiennent un schwa potentiel (cf. Figure 6-85) :

- soit interne : les cas de *petit* et *petits* ;
- soit final : les cas de *sommes*, *grandes* et *êtes* ;
- soit à la fois un schwa interne et un schwa final : le cas de *petites*.

Position du schwa	Forme	Monosyllabique	Polysyllabique
Schwa interne	<i>petits</i>	[pti(z)]	[pəti(z)]
	<i>petit</i>	[pti(t)]	[pəti(t)]
Schwa final	<i>sommes</i>	[sɔm(z)]	[sɔmə(z)]
	<i>grandes</i>	[gʁãd(z)]	[gʁãdə(z)]
	<i>êtes</i>	[et(z)]	[etə(z)]
Schwa interne et final	<i>petites</i>	[pətiti(z)]	[pətitiə(z)]

Figure 6-85 Formes polysyllabique et monosyllabique selon la (non)-réalisation du schwa

Une analyse de l'influence de la longueur sur la réalisation de la liaison nécessite donc une vérification de la longueur de certaines des formes liantes, vérification qui ne peut se satisfaire d'une simple observation graphique.

3. Le problème de la forme

La variation observée, par exemple à l'intérieur du paradigme flexionnel de *être* et de *avoir*, montre que la liaison est un problème qui se pose pour la forme en emploi. Cela n'aurait pas de sens de discuter de *être* ou de *avoir*, mais aussi de *petit* ou de *grand* de façon générale. Il faut donc procéder forme par forme. Cependant, procéder forme par forme ne signifie pas qu'il ne faut pas tenir compte du lien forme / lexème, car une même forme peut-être associée à deux lexèmes différents. C'est le cas par exemple de *en* ou de *sommes*.

4. Le problème de la catégorie

Tenir compte du lien forme / lexème signifie *de facto* tenir compte de l'information catégorielle (cf. Milner, 1989 sur les propriétés identificatoires des unités lexicales).

C'est le cas de *en* ou de *sommes* cités plus haut, respectivement pronom / préposition et nom / verbe. Mais aussi de *petits* et de *grands*, qui peuvent être adjectif ou nom ; de *son* qui peut être nom ou possessif, de *bien*, adverbe ou nom, *etc.*

L'existence de ces formes polycatégoriques (par ex. *en*) implique que selon la catégorie employée (préposition ou pronom dans le cas de *en*), le discours sur la liaison sera différent.

5. Le problème de la syntaxe et de la cohésion syntaxique

Ce problème se manifeste de deux manières différentes :

1. d'une part, on observe de fortes variations selon le contexte syntaxique. La cohésion syntaxique entre un Spécifieur de N et le N (par ex. *mon_ami*) est manifestement beaucoup plus forte qu'entre un V et son argument (par ex. *avait_un_ami*). Il faudra donc tenir compte du type de contextes dans lequel la liaison opère ;

2. d'autre part, un mot pourvu d'une catégorie en emploi ne va pas avoir le même fonctionnement qu'un autre mot de la même catégorie dans un même emploi. On note par exemple, que *pas* et *plus* adverbes, n'ont pas les mêmes taux de réalisation de la liaison : respectivement 1,4 % de LR et 64,1 %.

6. Le problème du lexique

Observer la fréquence des CL mot par mot est certes un artefact mais permet tout de même de voir que certains mots liaisonnent plus que d'autres.

Ce qui reste à déterminer, c'est la fréquence de ces occurrences. Par exemple, les formes à final /z/ sont très nombreuses, mais finalement peu de liaisons sont réalisées, sauf pour certaines formes : il faut alors vérifier que les moins fréquentes ne réalisent pas la liaison.

Par ailleurs, la fréquence de réalisation de la liaison peut être très différente pour des formes ayant le même nombre de contextes (mais pas la même CL). Nous proposons ci-dessous deux exemples illustrant nos propos :

	Nombre total dans <i>EI</i>	% de Liaisons Réalisées
<i>son</i>	58	100 %
<i>vais</i>	51	3,92 %
<i>quand</i>	639	77,9 %
<i>puis</i>	639	0,16 %

Figure 6-86 Exemples de formes ayant le même nombre de contextes mais des taux de LR très distincts

Ce constat amène à s'interroger sur la catégorie grammaticale à laquelle appartiennent ces unités. Cette question semble importante quand on voit que pour /n/, il y a peu de catégories en jeu (et une suspecte puisque la liaison n'est pas réalisée à 100%) et qu'il semble y avoir beaucoup de catégories en jeu pour /t/ et /z/.

Il apparaît donc qu'une analyse en fonction de la CL ne permet pas d'expliquer le fonctionnement de la liaison. Cependant, elle permet de faire émerger un certain nombre de problèmes qui ont principalement trait à la longueur des formes, à leur emploi en contexte et à la cohésion syntaxique.

Dans les sections suivantes, nous revenons sur le problème de la longueur (*cf. ci-dessous*) et traiterons des aspects syntaxiques dans la section 6.5.

6.4 Influence de la longueur

L'analyse des consonnes de liaison /n/, /t/ et /z/ a permis de mettre en évidence un certain nombre de problèmes relatifs à la longueur.

Premièrement, l'examen des formes montre que (i) /n/ est la consonne de liaison la plus souvent réalisée (96,6 % de LR) et que (ii) tous les contextes liaisonnants identifiés (EILRn) sont monosyllabiques.

Deuxièmement, les CL /z/ et /t/ ne réalisent la liaison que dans 59,1 % et 32,4 % des contextes, respectivement, et nous avons observé qu'un certain nombre des formes avec une CL, en raison du schwa qu'elles contiennent, pouvaient aussi bien se réaliser de façon monosyllabique que de façon polysyllabique.

Dans le Chapitre 3 (*cf. 3.4.1, pp. 97 et sq.*), nous avons rappelé les 3 hypothèses qui s'étaient attachées à montrer l'influence de la longueur sur la réalisation de la liaison :

Hypothèse 1 : l'influence de la longueur de M1 ;

Hypothèse 2 : l'influence de la longueur du groupe syntagmatique contenant M2 ;

Hypothèse 3 : l'influence de la longueur du groupe syntagmatique contenant M1.

La différence entre les trois hypothèses porte en fait sur la nature de l'élément à prendre en compte : un mot ou un groupe syntagmatique. Mais elles s'accordent toutes pour associer une plus forte réalisation de la liaison à un « élément » court (qu'il s'agisse donc d'un mot ou d'un groupe syntagmatique).

Cette section s'attachera à faire le point sur l'influence possible de la longueur des éléments dans la réalisation de la liaison. Toutefois, pour ce faire, il est nécessaire de préciser plusieurs points :

1. Comme nous l'avons vu en 5.1.3 (pp. 160 et *sq.*), le codage liaison de PFC informe sur la longueur de M1 (monosyllabique ou polysyllabique), mais ne fournit aucune information sur la longueur de M2, ni sur la longueur du groupe syntagmatique auquel M1 appartient, ni sur la longueur du groupe syntagmatique auquel M2 appartient. Ce que nous avons qualifié de « problème de la granularité du codage » (cf. 5.1.3.3.3, p. 164) pose ici véritablement problème puisque nos observations ne pourront porter que sur la longueur de M1. Autrement dit, nous ne testerons que la première hypothèse (Hypothèse 1) qui demeure toutefois l'hypothèse la plus envisagée, et laisserons de côté les Hypothèses 2 et 3 ;

2. Dans la mesure où nous disposons des résultats par consonne de liaison, même si telle ou telle consonne n'a pas fait l'objet d'une hypothèse précise, nous présenterons les résultats par consonne de liaison. Nous rassemblerons nos observations dans le bilan ;

3. Rappelons que (i) nous avons identifié les contextes réalisés et non réalisés à partir d'un ensemble de formes ; (ii) le codage PFC distingue uniquement les contextes de liaisons erratiques non réalisées de tous les autres contextes (les catégoriques, les variables et les erratiques réalisées) ; (iii) le codage PFC n'a prévu aucun étiquetage morpho-syntaxique. De ce fait, nos calculs sur la longueur s'appuieront aussi bien sur des contextes catégoriques que des contextes variables, bien que l'influence de la longueur soit surtout intéressante à examiner dans le cas des contextes variables. Comme nous le verrons par la suite, le fait d'avoir considéré l'ensemble de ces contextes n'a que peu d'incidence sur la validité de l'hypothèse testée ;

4. Enfin, concernant /z/ et /t/, du fait que la longueur de certaines des formes réalisées peut varier, il apparaît nécessaire dans un premier temps, d'identifier les contextes ambigus : il s'agira de préciser exactement si ces formes sont monosyllabiques ou polysyllabiques, pour pouvoir, par la suite, nous concentrer sur la réalisation de la liaison de chacune de ces formes.

6.4.1 Répartition des monosyllabes et des polysyllabes

L'ensemble des contextes identifiés pour chaque EI se répartit de la façon suivante (cf. également Figure 6-26, p. 219) :

EI	Nombre de contextes	Nombre de liaisons réalisées	Nombre de liaisons non réalisées
/n/	4 366	4 229	137
/t/	4 972	1 610	3 362
/z/	8 437	4 984	3 453
Total	17 775	10 823	6 952

Figure 6-87 Nombre de contextes pour les CL /z/, /n/ et /t/

Sur l'ensemble de toutes les formes identifiées, 16 173 sont monosyllabiques et 1 602 sont polysyllabiques. La répartition pour chacune des CL est la suivante :

- toutes les formes avec /n/ sont monosyllabiques, soient 4 366 formes ;
- 3 660 formes avec /t/ sont monosyllabiques et 1 312 sont polysyllabiques ;
- 8 147 formes avec /z/ sont monosyllabiques et 290 formes sont polysyllabiques¹²².

Ces données sont représentées dans le tableau suivant :

	Contextes monosyllabiques		Contextes polysyllabiques		Total
	Nombre	%	Nombre	%	
/n/	4 366	100 %	0	0 %	4 366
/t/	3 660	73,61 %	1 312	26,39 %	4 972
/z/	8 147	96,56 %	290	3,44 %	8 437
Total	16 173	90,99 %	1 602	9,01 %	17 775

Figure 6-88 Répartition des contextes de liaison en /n/, /t/ et /z/ selon la longueur

Ce tableau indique donc que la plupart des contextes de liaison sont monosyllabiques. Ils représentent en effet près de 91 % de l'ensemble des contextes

¹²² Si l'on considère que les déterminants et les pronoms de notre corpus ne se seraient réalisés qu'en contextes catégoriques (ce que nous n'avons donc pas eu les moyens de vérifier), la différence de réalisation entre les monosyllabiques et les polysyllabiques est de l'ordre de 3 %. Ce problème méthodologique ne se pose que pour la consonne /z/.

identifiés. Ce constat est valable pour l'ensemble des CL : toutes les formes en /n/ sont monosyllabiques, 73,6 % des formes en /t/ sont monosyllabiques et près de 96,6 % des formes en /z/ sont monosyllabiques.

Concernant plus précisément les CL /t/ et /z/, un certain nombre de formes présentent deux possibilités de réalisation. Du fait qu'elles contiennent toutes un schwa, soit interne, soit final (cf. 6.3.3.3), elles peuvent être soit monosyllabiques, soit polysyllabiques.

1. La CL /t/

Sur les 48 contextes relevés pour la forme *petit*, 12 sont monosyllabiques et 36 sont polysyllabiques.

Aussi, la répartition de l'ensemble des formes monosyllabiques et polysyllabiques est-elle de la forme suivante :

Monosyllabes		Polysyllabes	
Forme	Nombre de contextes	Forme	Nombre de contextes
<i>c'est</i>	1 470	<i>avaient</i>	408
<i>quand</i>	639	<i>avait</i>	374
<i>est</i>	636	<i>était</i>	343
<i>tout</i>	287	<i>étaient</i>	71
<i>sont</i>	208	<i>allait</i>	45
<i>ont</i>	178	<i>petit</i>	36
<i>faut</i>	127	<i>fallait</i>	35
<i>peut</i>	94		
<i>petit</i>	12		
<i>grand</i>	9		
Total	3 660		1 312

Figure 6-89 Répartition des formes monosyllabiques et polysyllabiques avec la CL /t/

Le tableau ci-dessus indique qu'il y a 10 formes monosyllabiques et 7 formes polysyllabiques contenant la CL /t/.

2. La CL /z/

Cinq formes ont été identifiées comme pouvant présenter une réalisation monosyllabique ou polysyllabique : *sommes*, *petits*, *grandes*, *petites* et *êtes*. Le tableau ci-dessous présente leur répartition en fonction de leur longueur :

Forme	Monosyllabique	Polysyllabique	Total
<i>petits</i>	8	14	22
<i>grandes</i>	6	3	9
<i>sommes</i>	15	12	27
<i>petites</i>	3	10	13
<i>êtes</i>	13	3	16

Figure 6-90 **Formes monosyllabiques ou polysyllabiques à finale /z/**

La répartition des différentes formes relevées selon leur longueur est représentée dans le tableau ci-dessous :

Monosyllabes		Polysyllabes	
Forme	Nombre de contextes	Forme	Nombre de contextes
<i>les</i>	1 047	<i>toujours</i>	184
<i>pas</i>	880	<i>assez</i>	41
<i>ils</i>	864	<i>avons</i>	24
<i>des</i>	796	<i>petits</i>	14
<i>mais</i>	791	<i>sommes</i>	12
<i>puis</i>	639	<i>petites</i>	10
<i>vous</i>	433	<i>êtes</i>	3
<i>suis</i>	430	<i>grandes</i>	2
<i>nous</i>	427		
<i>dans</i>	398		
<i>deux</i>	334		
<i>plus</i>	248		
<i>trois</i>	211		
<i>mes</i>	148		
<i>très</i>	145		
<i>ans</i>	130		
<i>chez</i>	64		
<i>ses</i>	62		
<i>vais</i>	51		
<i>sommes</i>	15		
<i>êtes</i>	13		
<i>petits</i>	8		
<i>grands</i>	7		
<i>grandes</i>	3		
<i>petites</i>	3		
Total	8 147		290

Figure 6-91 Répartition des formes monosyllabiques et polysyllabiques avec la CL /z/

On observe qu'il y a 25 formes monosyllabiques et 8 formes polysyllabiques contenant la CL /z/.

Etant donné que nous avons déjà établi que les formes avec la CL /n/ sont toutes monosyllabiques et que la liaison est réalisée dans 96,9 % des contextes identifiés (cf. Figure 6-30, p. 221), l'examen détaillé des liaisons avec monosyllabes ou polysyllabes ne portera que sur les formes dont la CL réalisable est /t/ ou /z/.

6.4.2 La liaison avec les monosyllabes

Le nombre total de monosyllabes en contexte de liaison avec les CL /n/, /t/ et /z/ s'élève à 16 173 et la liaison est réalisée dans 66,3 % de ces contextes, comme le détaille le tableau ci-dessous :

	Nombre de contextes	Nombre de LR	% de LR
/n/	4 366	4 229	96,86 %
/t/	3 660	1 530	41,80 %
/z/	8 147	4 957	60,84 %
Total	16 173	10 716	66,25 %

Figure 6-92 La liaison avec les monosyllabes

Les taux de réalisation de la liaison sont relativement élevés pour chacune des CL. Seules les liaisons avec monosyllabes en /t/ sont réalisées en-dessous de 50 % (41,8 %).

Dans le tableau ci-dessous, nous avons regroupé l'ensemble des formes monosyllabiques en /t/ et les avons classées selon les taux de réalisation de la liaison relevés et ce, de façon décroissante¹²³ :

M1	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
<i>petit</i>	12	11	1	91,67 %
<i>tout</i>	287	238	49	82,93 %
<i>quand</i>	639	498	141	77,93 %
<i>grand</i>	9	5	4	55,56 %
<i>est</i>	636	279	357	43,87 %
<i>peut</i>	94	29	65	30,85 %
<i>c'est</i>	1 470	413	1 057	28,10 %
<i>sont</i>	208	40	168	19,23 %
<i>ont</i>	178	17	161	9,55 %
<i>faut</i>	127	0	127	0 %
Total	3 660	1 530	2 130	41,80 %

Figure 6-93 Les formes monosyllabiques liaisonnant en /t/

¹²³ Ce procédé sera répété pour le tableau regroupant les formes monosyllabiques en /z/ et les tableaux regroupant les formes polysyllabiques en /t/ et en /z/ (cf. 6.4.3.).

On observe que le nombre de contextes ne favorise pas la réalisation de la liaison. Par exemple, *c'est*, présent dans 1 470 contextes, ne réalise la liaison que dans 28,1 % des cas, alors qu'avec *peut*, pour lequel on ne compte que 94 contextes, la liaison est réalisée dans 30,9 % des contextes.

Dans le tableau ci-dessous, nous avons regroupé l'ensemble des formes monosyllabiques en /z/ :

M1	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
<i>grandes</i>	3	3	0	100 %
<i>petites</i>	3	3	0	100 %
<i>deux</i>	334	334	0	100 %
<i>mes</i>	148	148	0	100 %
<i>ses</i>	62	62	0	100 %
<i>ils</i>	864	853	11	98,73 %
<i>les</i>	1 047	1 030	17	98,38 %
<i>vous</i>	433	425	8	98,15 %
<i>des</i>	796	769	27	96,61 %
<i>très</i>	145	140	5	96,55 %
<i>dans</i>	398	378	20	94,97 %
<i>trois</i>	211	182	29	86,26 %
<i>chez</i>	64	53	11	82,81 %
<i>nous</i>	427	320	107	74,94 %
<i>grands</i>	7	5	2	71,43 %
<i>plus</i>	248	159	89	64,11 %
<i>sommes</i>	15	8	7	53,33 %
<i>petits</i>	8	4	4	50 %
<i>êtes</i>	13	4	9	30,77 %
<i>suis</i>	430	58	372	13,49 %
<i>vais</i>	51	2	49	3,92 %
<i>pas</i>	880	12	868	1,36 %
<i>mais</i>	791	4	787	0,51 %
<i>puis</i>	639	1	638	0,16 %
<i>ans</i>	130	0	130	0 %
Total	8 147	4 957	3 190	60,84 %

Figure 6-94 Les formes monosyllabiques liaisonnant en /z/

De la même façon que pour les formes monosyllabiques liaisonnant en /t/, les données montrent que le nombre de contextes ne prédit pas, *a priori*, une liaison plus souvent réalisée.

Par exemple, *puis* est présent dans 639 contextes et ne réalise la liaison que dans 0,16 % des cas, alors que *très*, que l'on retrouve dans seulement dans 145 contextes, réalise la liaison à près de 96,6 %. Comme autre exemple illustratif, l'examen des formes *ils* et *pas* indique que, indépendamment de tout autre facteur, pour un nombre quasi identique de contextes, 864 et 880 respectivement, les taux de réalisation de la liaison sont diamétralement opposés : la liaison avec *ils* est réalisée dans près de 99 % des contextes, alors qu'avec *pas*, elle est pratiquement inexistante (environ 1,4 % de réalisation).

6.4.3 La liaison avec les polysyllabes

Le nombre total de polysyllabes en contextes de liaison avec les CL /n/, /t/ et /z/ est de 1 602, et la liaison est réalisée dans seulement 7,4 % des contextes. Le tableau ci-dessous donne le détail de ces réalisations pour chacune des trois CL considérées :

	Nombre de contextes	Nombre de LR	% de LR
/n/	0	0	/
/t/	1 312	80	6,09 %
/z/	290	39	13,45 %
Total	1 602	119	7,43 %

Figure 6-95 La liaison avec les polysyllabes

Le tableau suivant regroupe l'ensemble des formes polysyllabiques en /t/ :

M1	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
<i>petit</i>	36	25	11	69,44 %
<i>étaient</i>	71	8	63	11,27 %
<i>était</i>	343	29	314	8,45 %
<i>allait</i>	45	2	43	4,44 %
<i>avait</i>	374	10	364	2,67 %
<i>avaient</i>	408	6	402	1,47 %
<i>fallait</i>	35	0	35	0 %
Total	1 312	80	1 232	6,10 %

Figure 6-96 Les formes polysyllabiques liaisonnant en /t/

Les formes polysyllabiques liaisonnants en /t/ sont peu nombreuses : 7 au total. On remarque que les taux de réalisation de la liaison sont, pour la plupart d'entre eux, faibles (entre 0 et 11,3 %) et ce, indépendamment du nombre de contextes. Par exemple, *avaient* se retrouve dans 408 contextes mais ne réalise la liaison que dans 1,5 % d'entre eux, alors que *étaient*, présent dans seulement 71 contextes, réalise la liaison à hauteur de 11,3 %.

Seul le pourcentage de réalisation de la liaison avec *petit* est élevé (69,4 %), bien que le nombre de contextes soient assez restreint (36 au total) et par ailleurs quasi équivalent à celui de la seule forme qui ne réalise aucune liaison : *fallait* (35 contextes, 0 réalisation).

Le tableau ci-dessous regroupe l'ensemble des formes polysyllabiques en /z/ :

M1	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
<i>sommes</i>	12	12	0	100 %
<i>petites</i>	10	9	1	90 %
<i>petits</i>	14	9	5	64,29 %
<i>grandes</i>	2	1	1	50 %
<i>êtes</i>	3	1	2	33,33 %
<i>avons</i>	24	5	19	20,83 %
<i>assez</i>	41	2	39	4,88 %
<i>toujours</i>	184	0	184	0 %
Total	290	39	251	13,45 %

Figure 6-97 Les formes polysyllabiques liaisonnant en /z/

Concernant les liaisons en /z/ avec un polysyllabique, on remarque que (i) la forme présente dans le plus de contextes (*toujours*) est également la forme qui ne réalise aucune liaison (0 %) et que (ii) cette forme est également la seule à présenter une consonne fixe précédant la CL et pouvant s'enchaîner avec M2. D'autres formes peuvent également présenter un enchaînement avec M2 dans le cas où la liaison n'est pas réalisée (*sommes*, *petites*, *grandes* et *êtes*). Mais ce qui les différencie d'avec *toujours* est le fait qu'elles contiennent toutes un schwa en final. Leurs taux de réalisation de la liaison sont d'ailleurs très hétérogènes et le nombre de contextes semble, encore une fois, n'avoir aucune incidence sur le nombre de liaisons réalisées : *assez* est présent dans 41 contextes mais la liaison n'est réalisée que dans 4,9 % d'entre eux, alors que *petits* que l'on retrouve dans 14 contextes, réalise la liaison à hauteur de 64,3 %.

6.4.4 Bilan

Dans cette section, nous avons testé l'hypothèse selon laquelle la longueur de M1 influence la réalisation de la liaison. Avec les outils à notre disposition et selon la méthodologie que nous avons adoptée, il apparaît que l'hypothèse est vérifiée par les données PFC : la liaison y est plus souvent réalisée quand on a un monosyllabe que lorsque l'on a un polysyllabe : les monosyllabes lient dans 66,3 % des contextes, les

polysyllabes dans 7,4 % des contextes. Toutefois, les deux autres hypothèses alternatives n'ont pas pu être examinées.

Par ailleurs, nous avons présenté les résultats par consonne de liaison, et les différences observées par consonne de liaison, tout au moins pour /t/ et /z/, ne semblent pas significatives. Aussi, si l'on pouvait poursuivre l'analyse de l'influence de la longueur, il semblerait pertinent de distinguer les CL.

6.5 Syntaxe

L'objectif de cette partie, qui concerne plus particulièrement l'analyse des réalisations de la liaison sous leur aspect syntaxique, consiste à reprendre la tripartition de Delattre (1947), révisée par les travaux ultérieurs, et à la confronter à notre propre corpus.

Ces propositions de départ sont résumées dans le tableau suivant¹²⁴ :

¹²⁴ On se reportera aux tableaux détaillés des nouvelles répartitions des contextes de liaisons erratiques, catégoriques et variables, présentés en bilan de la section concernant la tripartition des contextes de liaison (*cf.* 3.2.3.2.4, pp. 76 et *sq.*).

	M1	M2
Liaisons erratiques	nom singulier	
	nom propre	
	pronom (pas personnel) nasal	
	pronom personnel postposé nasal 3ème personne	- pronom personnel sujets (<i>on, ils, elles</i>) - pronom complément (<i>en, les</i>)
	adverbe interrogatif	
	conjonction polysyllabique	
	conjonction <i>et</i>	
		<i>h aspiré</i>
		<i>oui, non</i>
	- adjectif singulier - pronom nasal	invariable
		- un, huit, onze - les dérivés : <i>unième, huitième, onzième</i>
	nom composé au pluriel	
	groupe figé sans liaison	
nom pluriel	verbe	
Liaisons catégoriques	article	nom
	adjectif	nom
	déterminatif	- pronom - adjectif
	pronom personnel	- verbe - pronom + verbe
	verbe	pronom personnel
	pronom personnel postposé au verbe	pronom
	préposition <i>en</i>	
	groupes figés	
Liaisons variables		invariable
	nom pluriel	- adjectif - invariable
	pronom non personnel	
	verbe	
		participe passé
	préposition polysyllabique	
	adverbe polysyllabique	
	conjonction monosyllabique sauf <i>et</i>	
	adverbe monosyllabique	
	préposition monosyllabique sauf <i>en</i>	
	adjectif qualificatif pluriel	
	adjectif qualificatif singulier sauf <i>petit</i>	

Figure 6-98 Tableau révisé de la grille tripartite de Delattre

Pour différentes raisons que nous exposons ci-dessous, nous ne pourrions tester tous les contextes de la grille.

Comme nous l'avons vu dans la section 5.1.3 (pp. 160 et *sq.*), concernant le codage liaison, PFC opère des distinctions uniquement entre les contextes de liaison erratique non réalisée, qu'il ne code pas, et tous les autres contextes qui sont, eux, codés (liaisons erratiques réalisées, liaisons catégoriques et liaisons variables). De ce fait, nous excluons d'ores et déjà de notre champ d'investigation, tous les contextes de liaison erratique¹²⁵, à l'exception du contexte « N pluriel + Verbe ». En effet, on n'a proposé que très récemment son intégration dans la classe des liaisons erratiques (*cf.* 3.2.3.2.3). Comme dans la grille tripartite de Delattre (1947), il comptait parmi les liaisons variables, tous les contextes « N pluriel + Verbe » ont été codés par PFC.

Nous présentons ci-dessous d'autres contextes, faisant cette fois-ci partie de la classe des liaisons catégoriques et / ou variables, que nous excluons également :

1. Les contextes de liaison pour lesquels les deux catégories (M1 et M2) sont spécifiées

Le codage liaison prévu par PFC ne porte que sur un seul des éléments en jeu dans le processus de liaison, soit M1. Comme aucun étiquetage catégoriel n'est disponible, nos tests ne peuvent porter que sur les contextes dans lesquels une seule catégorie est exprimée (par ex. : « Verbe + »). Les contextes contenant deux catégories¹²⁶ sont donc éliminés en raison de notre impossibilité à identifier la catégorie de M2 (en tout cas de manière générale et non au cas par cas).

2. Les contextes dans lesquels seul M2 est spécifié

Conséquemment, étant donné que les seules informations disponibles portent sur le contexte gauche, nous excluons également les contextes dans lesquels c'est la catégorie de droite qui est fournie¹²⁷.

¹²⁵ *Cf.* Annexe 13 pour la liste complète des contextes de liaison erratique.

¹²⁶ *Cf.* Annexe 14 pour la liste complète des contextes pour lesquels les deux catégories (M1 et M2) sont spécifiées.

¹²⁷ *Cf.* Annexe 15 pour la liste complète des contextes dans lesquels seul M2 est spécifié.

3. Les contextes exprimés à l'aide de catégories formant une classe ouverte

Ensuite, on remarque que les catégories proposées par Delattre, se répartissent en classes fermées et en classes ouvertes.

Si l'identification de l'ensemble des lexèmes contenus dans l'espace des classes fermées est possible, il n'en va pas de même pour ceux qui relèvent d'une classe ouverte. En effet, comment pourrait-on mener une réflexion générale autour des adjectifs, dont le nombre total de formes relevées s'élève seulement à quatre dans notre corpus ? De la même façon, même si le nombre de formes verbales identifiées est de 16, sachant qu'il existe plus de 8 000 verbes en français et un nombre bien plus considérable de formes dès qu'elles sont fléchies, il est impossible de tirer une quelconque conclusion sur le comportement des « Verbe + ». Dans ces conditions, on ne peut traiter que des contextes exprimés à l'aide de catégories relevant de classes fermées. Tous les contextes exprimés à l'aide de catégories relevant d'une classe ouverte sont donc exclus¹²⁸.

La seule exception que nous ferons concerne les formes des temps simples des verbes *être* et *avoir*, (i) qui ont fait l'objet d'une attention particulière, notamment dans le rapport entre fréquence d'apparition de la liaison et la fréquence de la forme en emploi (cf. entre autres Ågren, 1973 ; Bybee, 2001) ; et (ii) qui sont très fréquents en français du fait qu'ils sont très souvent employés comme auxiliaires.

4. Les contextes pour lesquels il est impossible d'identifier la catégorie de la forme en contexte

Une fois établi que les catégories observables doivent faire partie de contextes dans lesquels seul le contexte gauche est spécifié, et que cette catégorie relève d'une classe fermée, encore faut-il pouvoir répondre, dans certains cas, à deux contraintes d'ordre catégoriel :

(i) connaître la catégorie de la forme observée en emploi. On peut, par exemple, mentionner *les*, qui peut aussi bien être déterminant (par ex. *les enfants*) que pronom (par ex. *donne-les à Gaby*) ;

(ii) et par là-même, connaître la position de cette forme dans la phrase. Par exemple, pour un pronom sujet tel que *vous*, il est impératif de savoir s'il est préposé

¹²⁸ Cf. Annexe 16 pour la liste complète des contextes exprimés à l'aide de catégories relevant d'une classe ouverte.

(par ex. *vous avez*) ou postposé au verbe (par ex. *avez-vous encore du travail*), le type de liaison proposé étant différent selon chacun des cas (catégorique pour le premier exemple ; variable pour le second).

Etant donné, encore une fois, qu'aucun étiquetage catégoriel n'a été prévu par PFC, cette vérification n'a pu se faire que si le nombre de contextes n'était pas trop élevé.

Devant l'étendue des difficultés rencontrées (manque d'information catégorielle, manque d'information sur l'ordonnement des catégories, indistinctions catégorielle, sémantique), nous nous sommes essentiellement attachée aux contextes suivants : « Conjonction monosyllabique + », « Préposition monosyllabique + », « Préposition polysyllabique + », « Adverbe monosyllabique + » et les formes fléchies des verbes *être* et *avoir* aux temps simples, soit une partie des contextes qui ont été révisés par des travaux postérieurs à Delattre :

Contextes considérés
conjonction monosyllabique sauf <i>et</i> + ¹²⁹ préposition monosyllabique + préposition polysyllabique + adverbe monosyllabique + temps simples de <i>être</i> et <i>avoir</i> +

Figure 6-99 Contextes de liaison considérés dans l'analyse syntaxique

En dehors des formes aux temps simples de *être* et *avoir*, l'ensemble des contextes examinés sont exprimés à l'aide de catégories relevant d'une classe fermée. Aussi, lorsque cela était nécessaire et possible, nous avons établi la liste des formes qui nous manquaient, puis nous avons effectué une nouvelle recherche de contextes¹³⁰, et ce, afin de pouvoir considérer ces catégories dans leur intégralité.

¹²⁹ Le contexte « *et* + » est décrit comme erratique (cf. Figure 3-12)

¹³⁰ Ces nouveaux contextes ne correspondent qu'à des formes qui peuvent présenter une liaison (i.e. contenant une CL) et qui ne sont pas indistinctes catégoriellement.

6.5.1 Conjonction monosyllabique +

Les contextes de liaison « Conjonction monosyllabique + » sont décrits comme des contextes de liaison variable.

Dans le tableau ci-dessous, nous présentons la liste complète des conjonctions monosyllabiques pouvant réaliser la liaison en français. Ces conjonctions, au nombre de deux, sont toutes deux présentes dans notre corpus et le nombre d'occurrences de chacune d'entre elles est de 639 pour *puis* et 791 pour *mais*.

Pour chacune d'entre elles, nous avons indiqué le nombre total de contextes dans lesquels elles sont présentes, ainsi que leur taux respectif de réalisation de la liaison. Par ailleurs, nous les avons classées par ordre alphabétique¹³¹ :

Conjonction monosyllabique	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
<i>mais</i>	791	4	787	0,51 %
<i>puis</i>	639	1	638	0,16 %
Total	1 430	5	1 425	0,35 %

Figure 6-100 Taux de LR en corpus des conjonctions monosyllabiques considérées

La liaison avec les conjonctions monosyllabiques est variable, voire quasi inexistante : 0,35 % de LR¹³². Cependant, on note qu'il existe des liaisons dans des

¹³¹ Cette façon de procéder sera utilisée pour chacun des contextes étudiés.

¹³² Delattre (1947) classe *quand* parmi les adverbes. Or, Grevisse (1969) le classe à la fois dans les adverbes de temps et les conjonctions de subordination. Nous n'avons donc pas intégré *quand* aux conjonctions monosyllabiques examinées, parce que (i) ce n'est pas notre objet que de discuter de l'éventuelle catégorie adverbiale de *quand*, et (ii) que, de toute façon, il nous est impossible de distinguer les *quand* conjonctions des *quand* adverbes (si tant est que cette catégorie est pertinente). Malgré cette difficulté identificatoire, on peut donner le résultat global pour la forme : sur 639 contextes, 498 réalisent la liaison, soient 77,9 %. Ces résultats sont concordants avec ceux de Ranson (2008), qui note que *quand* liaisonne à hauteur de 69 %. Par ailleurs, les résultats d'Encrevé (1988) indiquent qu'avec *quand*, la liaison est toujours réalisée même si, dans certains cas, elle est non enchaînée. Dans le corpus PFC, la forme *quand* présente clairement une liaison variable. Si l'on fait l'hypothèse que *quand* est majoritairement conjonction (ce qui semble avéré après un rapide examen des contextes), ce résultat pose problème, puisque d'un côté, *mais* et *puis* sont pratiquement toujours sans liaison et de l'autre, *quand* manifeste une liaison très fréquente. On peut avoir les plus grands doutes sur la pertinence de calculer une moyenne de réalisations, car les « individus » pris en compte dans la moyenne ont des comportements très différents. Et par suite, avoir les plus grands doutes sur la pertinence de tirer une conclusion sur le fonctionnement des prépositions monosyllabiques. Ce qui apparaît plutôt, c'est que la conjonction de coordination ne liaisonne pas, à la différence de *quand* (conjonction de subordination). Mais cela reste une autre hypothèse à tester.

expressions telles que *mais_il est venu*, *puis_il est venu*. La liaison est donc variable mais non réalisée dans le corpus.

6.5.2 Préposition monosyllabique +

La liaison dans une séquence du type « Préposition monosyllabique + » est décrite comme variable.

Dans la classe des prépositions monosyllabiques en français, on en dénombre 9 avec une CL en finale et pouvant donc réaliser la liaison avec le mot qui suit : *chez*, *dans*, *dès*, *en*, *hors*, *près*, *sans*, *sous* et *vers*.

La forme *en* peut aussi bien être employée comme préposition (par ex. *en avant*) que comme pronom (par ex. *il en a bu*). Or, le corpus PFC contient 959 occurrences de *en* en contexte de liaison. Certes l'examen des contextes de liaison non réalisée peut nous fournir un certain nombre d'indications. Mais il nous a été impossible de distinguer entre préposition et pronom dans les contextes de liaison réalisés. De ce fait, nous excluons de nos observations générales la forme *en*, qui sera traitée indépendamment.

Le tableau suivant regroupe l'ensemble des prépositions considérées, dont le nombre s'élève désormais à 8 :

Préposition monosyllabique	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
<i>chez</i>	66	55	11	83,33 %
<i>dans</i>	398	378	20	94,97 %
<i>dès</i>	0	/	/	/
<i>hors</i>	0	/	/	/
<i>près</i>	0	/	/	/
<i>sans</i>	31	30	1	96,77 %
<i>sous</i>	3	3	0	100 %
<i>vers</i>	9	0	9	0 %
Total	507	466	41	91,91 %

Figure 6-101 Taux de LR en corpus des prépositions monosyllabiques considérées

Il apparaît tout d'abord que 3 des prépositions sont absentes du corpus : *dès*, *hors* et *près*. Ensuite, on remarque qu'à l'exception de *vers*, pour lequel la liaison n'est réalisée dans aucun des 9 contextes relevés, les taux de réalisation de la liaison avec les prépositions monosyllabiques sont élevés : ils sont supérieurs à 83 % et la liaison est réalisée dans les 3 contextes dans lesquels on retrouve *sous*¹³³.

Il semble donc que la liaison avec les prépositions monosyllabiques est très fréquemment réalisée (91,9 % de LR au total). Ces résultats s'accordent avec ceux de Malécot (1975) qui observe que la liaison avec les prépositions monosyllabiques n'est variable qu'à hauteur de 8 %.

Concernant maintenant *en*, nous proposons ci-dessous son taux de réalisation de la liaison :

Forme	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
<i>en</i>	959	939	20	97,9 %

Figure 6-102 Taux de LR en corpus de la forme *en*

Les 20 contextes dans lesquels la liaison n'est pas réalisée, sont des contextes où *en* est toujours employé comme préposition et jamais comme pronom.

Mais il demeure le problème de son identification catégorielle. On ne sait pas à combien de contextes rapporter ces 20 occurrences non réalisées. Si la liaison dans des contextes « Préposition + » avec *en* est variable, alors cette observation irait à l'encontre des résultats proposés par Durand et Lyche (2008) qui montrent que seule la liaison avec *en* préposition demeure catégorique, toutes les autres étant désormais variables.

¹³³ *sous* peut également être employé comme nom, mais les trois occurrences de notre corpus sont exclusivement prépositionnelles.

6.5.3 Préposition polysyllabique +

Les contextes de liaison « Préposition polysyllabique + » sont décrits comme manifestant une liaison variable.

En français, les prépositions polysyllabiques contenant une CL sont au nombre de 11 : *après*, *avant*, *concernant*, *depuis*¹³⁴, *devant*¹³⁵, *durant*, *envers*, *hormis*, *moyennant*, *nonobstant* et *pendant*.

Certaines prépositions de cette liste sont susceptibles d’être aussi un nom singulier. Cependant, le protocole PFC recommande de ne pas coder ces contextes sauf si la liaison est réalisée. Or, il s’avère, après vérification, que toutes sont employées comme prépositions, aucune comme nom. Nous les conservons donc dans l’ensemble analysé.

Le tableau ci-dessous présente cet ensemble :

Préposition polysyllabique	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
<i>après</i>	155	0	155	0 %
<i>avant</i>	11	0	11	0 %
<i>concernant</i>	0	/	/	/
<i>depuis</i>	22	0	22	0 %
<i>devant</i>	14	0	14	0 %
<i>durant</i>	0	/	/	/
<i>envers</i>	0	/	/	/
<i>hormis</i>	0	/	/	/
<i>moyennant</i>	0	/	/	/
<i>nonobstant</i>	0	/	/	/
<i>pendant</i>	44	0	44	0 %
Total	246	0	188	0 %

Figure 6-103 Taux de LR en corpus des prépositions polysyllabiques considérées

On remarque tout d’abord que sur un total possible de 11 prépositions, seules 5 sont présentes dans notre corpus : *après*, *avant*, *depuis*, *devant* et *pendant*.

¹³⁴ Toutes les occurrences de *devant* réalisent le schwa contenu dans la première syllabe. *Devant* est donc ici polysyllabique.

¹³⁵ *Id. cf.* ci-dessus.

Les résultats affichés par ce tableau, amènent à une conclusion rapide : la liaison avec les prépositions polysyllabiques n'est jamais réalisée. Cela suggère-t-il pour autant que le contexte de liaison « Préposition polysyllabique + » doit intégrer la classe des liaisons erratiques ? Probablement pas si l'on considère qu'il est tout à fait possible d'avoir en français une séquence du type « *après_avoir mangé* », qui réaliserait la liaison. Manifestement, on se trouve ici confronté à un problème de corpus et de style.

Ces résultats amènent toutefois à s'interroger sur la notion de « variation » et sur ce qu'elle intègre.

Nous laissons pour le moment cette question de côté, pour poursuivre nos analyses.

6.5.4 Adverbe monosyllabique +

Les contextes de liaison « Adverbe monosyllabique + » sont décrits par la littérature comme manifestant une liaison variable.

Dans la classe des adverbes monosyllabiques en français, on a : *bien, fort, pas, plus, tout* et *très*.

Nous conservons *bien, pas, plus* et *tout*, bien qu'ils présentent un certain nombre de problèmes :

- *bien* et *tout* peuvent être employés comme nom singulier (par ex. *le bien ; le tout*). Mais PFC n'a prévu aucun codage pour ces contextes sauf lorsque la liaison est réalisée ;

- *plus* peut également être employé comme nom (par ex. *c'est lui qui en a pris le plus*) mais dans ce cas, sa consonne finale est généralement fixe et *plus* n'est alors pas codé ;

- enfin, *pas* peut lui aussi avoir un emploi nominal (par ex. *il a le pas lourd*) mais, comme nous l'avons vu précédemment (cf. note de bas de page 121, p. 243), tous les contextes dans lesquels *pas* apparaît indiquent que c'est un adverbe et non un nom.

On obtient alors le tableau suivant :

Adverbe monosyllabique	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
<i>bien</i>	185	79	106	42,7 %
<i>fort</i>	0	/	/	/
<i>pas</i>	880	12	868	1,36 %
<i>plus</i>	248	159	89	64,11 %
<i>tout</i>	287	238	49	82,9 %
<i>très</i>	145	140	5	96,55 %
Total	1 745	628	1 117	35,99 %

Figure 6-104 Taux de LR en corpus des adverbes monosyllabiques considérés

Seul l’adverbe *fort* est absent du corpus. Les 5 formes restantes présentent des taux de réalisation de la liaison très disparates. Si Malécot (1975) observe que la liaison avec les adverbes monosyllabiques est désormais variable, *très* et *plus* ne réalisant la liaison qu’à hauteur de 63 % et *pas* à hauteur de 37 %, nos résultats indiquent des écarts beaucoup plus importants, puisque les taux s’échelonnent entre 1,4 % (pour *pas*) et 96,6 % (pour *très*). Autrement dit, selon la forme, la liaison est quasi inexistante, quasi catégorique, et seuls *bien*, *plus* et *tout* manifestent une liaison franchement variable (respectivement 43 %, 64 % et 84 %).

Il apparaît que l’on ne peut pas conserver le contexte général des « Adverbes monosyllabique + » dans la classe des liaisons catégoriques étant donné que la plupart des formes manifestent de la variation.

- Cependant, il s’avère difficile de conclure sur les « Adverbe monosyllabique + ». Ils ne relèvent pas clairement de la classe des liaisons catégoriques, mais les intégrer dans la classe des liaisons variables se révèle également compliqué du fait (i) qu’ils présentent des écarts de réalisation très importants les uns par rapport aux autres (1,4 % pour *pas*, 42 % pour *bien*, 83 % pour *tout*), (ii) le nombre d’occurrences n’est pas comparable (880 pour *pas* et 145 pour *très*).

- Si ces formes peuvent toutes être considérées comme des adverbes, elles ont, dans le même temps, des fonctionnements syntaxiques très différents (conduisant à des contextes de liaison qui ne manifestent pas forcément la même cohésion syntaxique). À côté de *bien évidemment il va y arriver ; c’est un garçon bien original*, on a *très honnêtement il va y arriver ; c’est un garçon très original*. Mais le parallèle n’existe pas entre *on va bien y arriver ; il se comporte bien* et **on va très y arriver ; *il se comporte*

très. La pertinence d'un traitement forme par forme doit alors être considérée. Nous y reviendrons dans notre conclusion générale.

6.5.5 Les temps simples de *être* + et *avoir* +

Les contextes de liaison « Verbe + » sont décrits comme présentant une liaison variable.

Nous ne traiterons ici que des temps simples des verbes *être* et *avoir*, soit les formes contenant une CL au présent de l'indicatif, à l'imparfait et au futur.

6.5.5.1 *être* +

Dans les formes fléchies de *être*, la forme *suis* peut aussi bien correspondre à une forme fléchie de *être* qu'à une forme de *suivre*. Le nombre total d'occurrences relevées s'élève à 430 (dont 58 réalisées) et il nous a été impossible d'établir avec exactitude de quelle forme il s'agissait. Aussi, nous excluons cette forme de nos observations.

Le nombre total de formes considérées est de 5 formes au présent de l'indicatif, 5 à l'imparfait et 4 au futur, soit un total de 14 formes, comme on peut le voir dans le tableau ci-dessous :

Temps	Formes	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
Présent de l'indicatif	<i>es</i>	53	0	53	0
	<i>est</i>	636	279	357	43,87 %
	<i>sommes</i>	27	12	15	44,44 %
	<i>êtes</i>	16	5	11	31,25 %
	<i>sont</i>	208	40	168	19,23 %
Imparfait	<i>étais</i>	12	1	11	8,33 %
	<i>était</i>	343	29	314	8,45 %
	<i>étions</i>	4	0	4	0 %
	<i>étiez</i>	2	0	2	0 %
	<i>étaient</i>	71	8	63	11,27 %
Futur	<i>seras</i>	0	/	/	/
	<i>serons</i>	0	/	/	/
	<i>serez</i>	0	/	/	/
	<i>seront</i>	2	0	2	0 %
	Total	1 374	374	1 000	27,22 %

Figure 6-105 Taux de LR en corpus des formes fléchies de *être* considérées

On remarque d'une part que les formes les plus représentées dans le corpus sont les 3^{ème} et 6^{ème} personnes du présent (*est* et *sont*), et la 3^{ème} personne de l'imparfait (*était*). D'autre part, les formes du futur sont très peu représentées puisque l'on ne compte que deux occurrences avec la 6^{ème} personne (*seront*), les autres formes étant absentes du corpus (*seras*, *serons* et *serez*).

De façon générale, la liaison avec *être* est variable et peu réalisée (27,22 %). Les taux de réalisation sont compris entre 8 % et 45 %. Les taux les plus faibles s'observent avec les formes de l'imparfait et les taux les plus forts, avec les formes du présent de l'indicatif. Seule la 6^{ème} personne (*sont*) accuse un taux de réalisation peu élevé (19,2 %) par rapport aux autres formes du même temps.

Ces résultats contrastent fortement avec les résultats d'Ågren (1973) qui indiquaient des taux de réalisation de la liaison beaucoup plus élevés que les nôtres. Par exemple, il notait que *est* réalisait la liaison à hauteur de 97 % alors que dans notre corpus, la liaison n'est faite que dans 43,9 % des cas. De la même façon, *sont* liaisonnait dans 86 % des contextes alors qu'ici, le taux n'est que de 19,2 %. Ågren explique les taux de réalisation de la liaison par la fréquence d'apparition de la forme en question : plus elle

est fréquente, plus la liaison est réalisée. Nos propres résultats ne peuvent cependant pas confirmer une telle hypothèse. En effet, les taux les plus forts s'observent dans des contextes dont le nombre d'occurrences est extrêmement différent.

Il en va de même pour les taux les plus faibles, comme on peut le voir dans le tableau suivant :

	Formes	Nombre de contextes	% de LR
Liaison +	<i>est</i>	636	43,87 %
	<i>sommes</i>	27	44,44 %
Liaison -	<i>était</i>	343	8,45 %
	<i>étais</i>	12	8,33 %

Figure 6-106 Comparaison de la fréquence de LR avec le taux de réalisation de la liaison avec les formes fléchies de *être*

Quant à la forme *suis*, nous fournissons seulement les résultats bruts :

Forme	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
<i>suis</i>	430	58	372	13,49 %

Figure 6-107 Taux de LR en corpus de la forme *suis*

6.5.5.2 *avoir* +

Concernant maintenant les formes fléchies de *avoir* au présent de l'indicatif, à l'imparfait et au futur, trois formes peuvent également être employées comme nom : *as*, *avions* et *auras*. Cependant, après vérification, il s'avère qu'aucune occurrence de *avions* et de *auras* ne sont des noms et que, dans le cas de *as*, si la forme relevait d'un emploi nominal, alors la consonne finale serait fixe, et la forme non codée. Aussi, ces trois formes seront étudiées au même titre que les autres formes qui ne présentent aucune ambiguïté.

L'ensemble de ces formes est présenté dans le tableau ci-dessous. Au total, on compte 4 formes au présent de l'indicatif, 5 à l'imparfait et 4 au futur, soit un total de 13 formes :

Temps	Formes	Nombre de contextes	Nombre de LR	Nombre de LNR	% de LR
Présent de l'indicatif	<i>as</i>	71	0	71	0 %
	<i>avons</i>	24	5	19	20,83 %
	<i>avez</i>	21	0	21	0 %
	<i>ont</i>	178	17	161	9,55 %
Imparfait	<i>avais</i>	10	0	10	0 %
	<i>avait</i>	374	10	364	2,67 %
	<i>avions</i>	4	0	4	0 %
	<i>aviez</i>	4	0	4	0 %
	<i>avaient</i>	408	6	402	1,47 %
Futur	<i>auras</i>	4	0	4	0 %
	<i>aurons</i>	0	/	/	/
	<i>aurez</i>	0	/	/	/
	<i>auront</i>	2	0	2	0 %
	Total	1 100	38	1 062	3,45 %

Figure 6-108 Taux de LR en corpus des formes fléchies de *avoir* considérées

Premièrement, ce tableau indique que deux des formes du futur ne font pas partie du corpus : *aurons* et *aurez*.

Ensuite, le nombre de liaisons réalisées est très faible : 38 sur 1 100, soit 3,5 % de réalisations seulement. Les taux de réalisation de la liaison sont tous très faibles (compris entre 0 % et 10 %), si l'on excepte *avons* qui réalise la liaison à hauteur de 20,8 %.

De la même façon que pour *être*, ces taux semblent indifférents au nombre de contextes possibles : *ont* est présent dans 178 contextes et réalisent la liaison dans 9,6 % des cas, alors que *avaient*, que l'on retrouve 408 fois, ne réalise la liaison que dans 1,5 % des cas (6 réalisations au total).

La liaison avec les formes fléchies de *avoir* aux temps simples est donc variable.

6.5.6 Bilan

L'objectif de cette partie consistait en la comparaison des contextes présents dans la grille tripartite de Delattre (avec ses révisions) avec les données que nous avons à notre disposition.

Avant toute observation générale sur les résultats obtenus, il apparaît nécessaire de faire un bilan sur certains problèmes que nous avons rencontrés.

Il semble tout d'abord que, de la même façon que pour l'analyse de l'influence de la longueur (*cf.* 6.4.4), les outils mis à notre disposition ne permettent pas une analyse fine des contextes syntaxiques, puisque le manque d'informations nous a très fortement limitée dans le nombre de contextes susceptibles d'être traités.

S'ajoute à ce point, le nombre très élevé de contextes au départ (plus de 17 000). La seule méthode à notre disposition a été d'identifier ces contextes à travers des formes / catégories discutées dans la littérature et formant une classe fermée. Nous ont échappé, bien évidemment, un nombre important de formes / catégories appartenant à une classe ouverte : les contextes de Delattre sont exprimés à l'aide de catégories relevant aussi de classes ouvertes, *de facto*, il nous était impossible de traiter de ces contextes. Au bilan, très peu ont pu être examinés.

Aussi, si l'on souhaite poursuivre l'analyse de la liaison à travers l'ensemble des contextes proposés par Delattre, il semblerait pertinent de se doter d'un étiqueteur morphosyntaxique qui permettrait alors de :

1. tester non seulement les contextes dans lesquels seul M1 est spécifié, mais de traiter également les contextes dans lesquels M1 et M2 sont spécifiés et ceux où seul M2 est spécifié. Le croisement des résultats obtenus permettrait de tester l'influence de la catégorie à laquelle appartient M2 ;

2. de tester les formes qui sont indistinctes d'un point de vue catégoriel, alors qu'en contexte, elles n'ont qu'une catégorie. Il est tout à fait dommageable d'éliminer un contexte de liaison très fortement représenté dans le corpus et par ailleurs largement discuté dans la littérature (*cf.* le cas de *en* pour ne donner qu'un exemple), sous prétexte que l'on ne peut « trier » les occurrences. Cela interroge par ailleurs, sur la pertinence des observations partagées, si personne n'a eu recours à ce genre d'outil.

Quant aux propositions de Delattre, révisées, elles sont exprimées à l'aide de catégories qui ne s'appuient sur aucune théorie syntaxique moderne. Un certain nombre de ces contextes sont exprimés de manière trop « simpliste » au regard de la complexité syntaxique des phrases du corpus (hésitations / reprises, expressions incidentes). De même, un certain nombre de catégories sont à revoir. Cependant, en reformulant les contextes de Delattre tout en tenant compte des travaux plus récents sur les catégories et sur la syntaxe, ce ne serait plus exactement les propositions de Delattre qui seraient testées¹³⁶. Le problème serait alors en partie décalé mais non résolu.

Aussi, si l'on veut avancer de manière précise sur la caractérisation de l'apport syntaxique dans la liaison, nous suggérons de prendre le problème tout à fait autrement, soit forme par forme, puis d'effectuer une analyse fine des contextes syntaxiques dans lesquels se trouve la forme, avant de tirer une quelconque généralisation.

Partir de la forme n'exclut nullement le recours à un savoir catégoriel et syntaxique. Mais celui-ci doit être fondé dans un rapport catégorie / organisation syntaxique. On peut faire l'hypothèse que cela se révélerait particulièrement utile pour nombre d'éléments relevant de classes fermées qui ont des fonctions syntaxiques très spécifiques.

Concernant maintenant les contextes que nous avons examinés : ils sont au nombre de 5 et (i) sont exprimés à l'aide de catégories relevant de classes fermées ; (ii) la catégorie prise en compte est celle de M1 ; enfin (iii) cette catégorie est suivie par un élément dont la catégorie n'est pas spécifiée (+).

1. De façon générale, ces contextes de liaison présentent tous des taux de réalisation avec des écarts plus ou moins importants ;
2. Les propositions présentées en 3.2.3.2.4 sont confirmées mais pas forcément avec les données du corpus (*cf.* a. et c.) :

b. les conjonctions monosyllabiques ne réalisent pratiquement jamais la liaison (0,35 % de LR sur 1 430 contextes), mais ce sont des liaisons variables ;

¹³⁶ En effet, tenir compte de travaux plus récents ne constitue pas un simple travail de réétiquetage (par ex. adjectif possessif / déterminant). C'est le rapport catégorie lexicale / organisation syntaxique qui est modifié (*cf.* Milner, 1989).

c. les prépositions monosyllabiques réalisent très fréquemment la liaison (LR > 83 %¹³⁷), mais ce sont des liaisons variables ;

d. les prépositions polysyllabiques ne réalisent jamais la liaison (0 % de LR sur un total de 246 occurrences), mais ce sont des liaisons variables ;

e. le cas des adverbes monosyllabiques est problématique : la variation observée s'échelonne entre 1,3 % (taux de LR avec *pas*) et 96 % (taux de LR avec *très*), ce qui incite à considérer que les formes doivent être étudiées une par une. Mais, globalement, ce sont des liaisons variables ;

f. enfin, les formes fléchies des verbes *être* et *avoir* aux temps simples présentent eux aussi une liaison variable. Avec *être*, les taux de LR sont de 27 % ; avec *avoir* de 3,5 %.

3. Les contextes a. et c. rappellent toute la justesse de l'observation de Delattre quant au fait qu'on ne fait de toute façon que très peu la liaison variable dans la conversation (*cf.* 6.2.4.2) ;

4. On touche par-là même aux discussions très actuelles sur la problématique de recueil de données et de leurs propriétés : opposition entre données attestées fournies par un corpus et données « inventées » soumises à un jugement (*cf.* Delais-Roussarie, 2008 pour des données en phonologie). En fait, cette opposition est complexe : les données attestées peuvent être grammaticales et agrammaticales (lapsus, erreurs de production) ; les données non attestées peuvent relever d'un problème de constitution de corpus ou de complétude, ou bien être non attestées parce que agrammaticales (*cf.* Cori et David, 2008).

¹³⁷ Nous ne considérons pas le taux de réalisation de la liaison avec *sous* pour lequel seuls 3 contextes ont été identifiés, ce qui nous semble trop peu pour pouvoir affirmer que cette préposition manifeste une liaison catégorique.

Chapitre 7 Conclusion générale

Dans le Chapitre 2, nous avons brièvement rappelé les différentes propositions qui ont cherché à caractériser le phénomène de la liaison. Nous avons ensuite décrit l'ensemble des dimensions qui interagissent dans la réalisation de la liaison et qui soulèvent un certain nombre de questions théoriques (*cf.* Chapitre 3). Le Chapitre 4 a été consacré exclusivement à la dimension phonologique. Le fait que beaucoup de travaux fondent leurs analyses sur la base empirique construite par leurs prédécesseurs, nous a conduit, avant tout choix théorique particulier, à faire le point sur les usages de la liaison dans le français contemporain. Pour cela, nous avons examiné les données proposées par le projet PFC (*cf.* 5.1). Ce projet, bien que présentant un certain nombre de problèmes récurrents à l'utilisation d'un corpus oral (*cf.* 5.1.3.3), a l'avantage de mettre à la disposition des chercheurs de nombreuses données contemporaines.

A l'issue de notre parcours, un bilan se dessine sur un certain nombre de points traités et il apparaît que les problèmes soulevés par le phénomène de la liaison demeurent, en partie, sans réponse, mais suggèrent de nouvelles pistes de recherche. Nous exposons ci-dessous l'ensemble de ces réflexions.

1. La dimension phonologique

Il apparaît tout d'abord que les liaisons non enchaînées, qui constituent le point central de l'ouvrage d'Encrevé (1988), sont en nombre trop restreint pour que l'on puisse les considérer comme pouvant intégrer une partie de la réflexion générale menée autour de la liaison : on ne dénombre que 130 contextes sur plus de 37 000 (cf. Figure 6-2). Il s'agit en fait d'un phénomène très particulier réservé aux spécialistes de la parole publique et donc à étudier de façon autonome.

Ensuite, l'hypothèse selon laquelle, en français, la liaison est réalisée pour éviter les hiatus et favoriser une syllabation ouverte, ne semble toujours pas pouvoir être confirmée. En effet, le français produit de nombreux hiatus, aussi bien internes qu'externes. De plus, dans de nombreux contextes, la liaison reste non réalisée alors même que sa réalisation permettrait l'enchaînement attendu. On ne peut donc conclure, en tout cas d'un point de vue descriptif, que le hiatus constitue en soi une motivation fondamentale de la liaison dans le français, comme l'ont si souvent décrit les normativistes (cf. Martinon, 1913 ; Brunot, 1966 par exemple). Si effectivement, on remarque la présence d'une consonne finale précédant la CL dans des contextes qui ne manifestent jamais la liaison (par ex. *toujours*), on ne peut manquer de noter que dans d'autres, la liaison n'est pas réalisée et qu'un hiatus est aisément produit (par ex. *je suis 10 allée*). D'autre part, nous avons mis en évidence que la question du hiatus ne se pose que dans les cas de liaisons variables et que les arguments en faveur d'une stratégie d'évitement du hiatus proviennent de l'examen des liaisons catégoriques (cf. 3.1.4). Aussi pensons-nous que, si l'on considère le phénomène de la liaison dans sa généralité, la question d'un évitement du hiatus ne constitue pas une des contraintes fortes régissant le comportement de la liaison.

Enfin, l'examen de la réalisation de la liaison à travers l'observation des consonnes de liaison a montré que :

(i) les CL /r/, /p/ et /k/ sont quantité négligeable. Elles ne présentent que très peu de contextes (1 042 contextes en tout), encore moins de formes possibles (essentiellement des formes verbales à l'infinitif pour /r/, deux formes *trop* et *beaucoup* pour /p/, et deux formes *long* et *franc* pour /k/), et ne se réalisent pratiquement jamais (1,3 % ELRr ; 11,2 % ELRp ; 0 % ELRk) ;

(ii) les CL nécessitant un examen particulier sont /n/, /t/ et /z/. Nous avons mis en évidence que /n/ est la CL qui présente le taux de réalisation le plus élevé (90,2 % ELRn) et le moins de formes en contexte de liaison (6 au total, toutes monosyllabiques), ce qui confirme les hypothèses présentées en 3.1.2. Les CL /t/ et /z/ présentent quant à elles des taux de réalisation beaucoup plus faibles, respectivement 23,27 % (dans ELRt) et 43,02 % (dans ELRz), mais sont présentes dans beaucoup plus de formes et de contextes syntaxiques. Cette analyse a donc mis en évidence que, du point de vue de leur taux de réalisation, leur ordre était /n/ > /z/ > /t/.

Cependant, cette observation n'a pas trouvé d'explication générale. On peut cependant émettre l'hypothèse que cet ordonnancement est lié à la multiplicité des dimensions contraignant la réalisation de la liaison, dimensions qui s'organisent différemment selon les contextes :

- /n/ est surtout réalisé dans des monosyllabiques dont le nombre est assez restreint mais pour lesquels les catégories syntaxiques sont très différentes (déterminant, pronom, préposition, adverbe, *etc.*). Aussi peut-on conclure que, ce qui réunit les formes avec une CL /n/ c'est leur longueur et non leur catégorie. Et que, de ce fait, la contrainte qui semble la plus forte est phonologique ;

- /z/ correspond essentiellement à une marque du pluriel et est présent dans de nombreux contextes, mais ces contextes ne manifestent pas forcément une cohésion syntaxique maximale. Aussi, elle est maximale entre un déterminant et un nom, entre un adjectif préposé et un nom, mais beaucoup moindre entre le sujet et le verbe.

- /t/ est surtout réalisé dans les contextes de liaison « Verbe + » dans lesquels la cohésion syntaxique est plus faible.

On voit que l'ordonnancement de ces contraintes est un problème complexe, qui demande bien d'autres études préalablement à toute généralisation. Nous y reviendrons dans le point 5 (multidimensionnalité et problème de formes).

2. L'interface phonologie-lexique

L'examen des formes contenant chacune des CL a permis d'analyser l'influence de la longueur et amène à la conclusion suivante : il apparaît que quelque chose se passe au niveau des monosyllabes, ce quelque chose semblant confirmer l'hypothèse 1 présentée en 3.4.1 et selon laquelle, plus M1 est court, plus la liaison tend à être réalisée.

Cependant, nous ne pouvons en apprécier les détails, puisque d'une part, cette hypothèse n'a pu être confrontée aux deux autres proposées dans la littérature et que, d'autre part, l'analyse a fait émerger l'importance de prendre en compte l'ensemble des dimensions en jeu. Il s'avère en effet que de nombreuses formes manifestent de la variation et ce, qu'elles soient monosyllabiques ou polysyllabiques. Par ailleurs, cette variation est plus ou moins importante selon la forme en question. D'un point de vue phono-lexical, il apparaît difficile de poser des contraintes générales puisque le phénomène semble devoir être caractérisé forme par forme.

3. La dimension syntaxique

La classification tripartite de Delattre et révisée par les travaux ultérieurs tente d'englober tous les contextes syntaxiques possibles de liaison. Cela permet de circonscrire les liaisons erratiques et catégoriques et d'observer que la catégorie des variables a considérablement augmenté, pour intégrer finalement beaucoup des contextes possibles.

Lorsqu'il s'est agi de reprendre le contenu de cette tripartition et de le soumettre à notre propre corpus, il s'est avéré impossible de tirer des généralités et ce, pour plusieurs raisons.

Tout d'abord, le corpus ne contenait pas l'ensemble des contextes proposés par la grille tripartite. Ensuite, pour les contextes présents, étant donné un certain nombre de difficultés auxquelles nous avons été confrontée (nombre élevé de contextes invérifiables « manuellement », manque d'outils spécifiques pour classer, extraire, etc. ; manque d'informations linguistiques), nous avons été choisis d'adopter une méthode d'analyse stricte : considérer uniquement les classes fermées (sauf pour *être et avoir*), les contextes dans lesquels seul M1 est spécifié et identifiable catégoriellement. Nous avons donc restreint le champ de nos observations à cinq contextes, tous décrits par la littérature comme réalisant la liaison de façon variable. Il est apparu que :

- (i) dans un contexte donné, les formes ont généralement des fonctionnements très différents ;
- (ii) la liaison ne touche jamais ou rarement la totalité d'une catégorie ;

(iii) enfin, on a des taux de réalisation très faibles, qui ne signifient pas pour autant que la liaison est erratique ou bien qu'elle n'est pas variable.

Cela implique que soit l'on a affaire à de simples tendances, soit qu'il faut revoir les catégories et les contextes proposés (cf. *très* et *bien*). On ne peut en tout cas pas tenir une généralisation sur toutes les *Prépositions* +, toutes les *Conjonctions* +, etc.

On pourrait s'étonner de ce manque de généralisations lorsque l'on considère l'ampleur du corpus de départ. Comment expliquer que, sur plus de 37 000 contextes codés dont près de 18 000 identifiés, l'on ne peut faire ressortir que si peu d'éléments de réponse aux différents aspects qui posent question ?

Mais ce que ce « problème » fait apparaître de manière positive est que (i) la liaison est un problème de formes en emploi ; (ii) il n'y a que très peu de formes susceptibles d'être impliquées dans le processus (cf. ci-dessous). Dès lors, pour les mettre au jour, il faut procéder à une analyse extrêmement fine sur une base comportant beaucoup d'éléments. De ce fait, nous revenons ici sur la nécessité de disposer d'outils permettant d'avoir accès, d'extraire et de trier des données sur la base de formes augmentées d'informations linguistiques (par ex. lemme en emploi (cf. la forme *suis*) ; catégorie en emploi, etc.).

Finalement, si les hypothèses catégorielles de Delattre sont justes, il apparaît que la syntaxe détermine dans la liaison somme toute peu de choses, puisque, si elle permet de distinguer les contextes de liaisons catégoriques et les contextes de liaison erratiques, les contextes de liaisons variables, qui représentent tout de même énormément des contextes possibles, sont laissés aux autres dimensions. Si, par contre, il faut revoir les catégories et la formulation des contextes, alors la part d'influence de la syntaxe reste à identifier.

4. La dimension sociolinguistique

Les résultats concernant la dimension sociolinguistique sont surprenants et vont à l'encontre de ce qui est généralement admis.

Tout d'abord, ils indiquent que, d'un point de vue stylistique, seule la distinction entre lecture de texte et conversation est pertinente, la différence entre la conversation guidée et la conversation libre restant infime. Ensuite, il s'est avéré qu'une comparaison

des taux de réalisation de la liaison selon le profil socioculturel des locuteurs ne permettait pas d'établir de « profil type » en ce sens, que deux locuteurs dont les parcours sont aux antipodes peuvent autant ou aussi peu réaliser la liaison. Seule la tranche d'âge semble pertinente puisqu'on s'aperçoit que la réalisation de la liaison est plus ou moins proportionnelle à la tranche d'âge : ce sont les locuteurs les plus âgés qui en réalisent le plus en situation conversationnelle et que cette différence générationnelle s'estompe avec la situation de lecture.

Ces résultats, en contredisant différentes études, ouvrent par là-même un nouveau chantier d'exploration.

5. Multidimensionnalité et problème de formes

Les travaux qui ont tenté de comprendre les mécanismes de la liaison, se sont principalement axés autour d'une seule voire de deux dimensions et ce, en raison du nombre de dimensions impliquées dans le processus. Ils ont chacun mis en évidence quelque chose. Cependant, les généralisations qu'ils proposent ne touchent qu'à leur point de vue et sont beaucoup trop fortes ; elles doivent faire face à de nombreux contre-exemples. En d'autres termes, les modèles qui cherchent à généraliser sur une base catégorielle, syntaxique, phonologique s'avèrent trop puissants.

En outre, ces modèles ne fournissent aucune prédiction pour la variation, justement parce qu'ils ne captent qu'une seule dimension. Et les rares travaux qui se sont intéressés à plus d'une dimension les envisagent le plus souvent une par une.

Un travail à l'interface des différents niveaux d'organisation de la langue, qui permettrait de déterminer l'articulation ou la hiérarchisation des contraintes qui seraient proposées, apparaît beaucoup plus prometteur.

Certes, on le sait depuis quelque temps déjà, et c'était même le projet de départ de cette thèse (*cf.* Chapitre 1). Les différents travaux sur la question ont même souligné la difficulté d'une telle entreprise ; leur position paraissait pessimiste devant la multiplicité des facteurs en jeu (*cf.* par exemple citation Ågren, p. 35). Mais finalement, on peut supposer que, du fait du petit nombre de formes en cause, cette entreprise qualifiée d'irréalisable par certains devienne possible. Ce que notre travail apporte, c'est la nécessité de travailler cette multidimensionnalité de manière fine, soit formes par formes.

Ce qui apparaît en effet, c'est une « pauvreté » quantitative du matériel empirique. Ce n'est pas parce qu'un mot est susceptible de liasonner qu'il est, dans le corpus, en contexte de liason. On a manifestement affaire à un « petit » ensemble de formes en contexte : 50 formes pour 17 775 contextes, représentant 64,7 % de l'ensemble des contextes considérés (cf. Figure 6-23 et Figure 6-27). Dans le cadre de l'analyse syntaxique, nous avons recherché les formes contenues dans les classes fermées contenues dans le corpus, et que nous n'avions pas identifiées : leur nombre s'élève à 9. Aussi, si l'on ajoute ces 9 formes aux 50 premières identifiées, on obtient un total de 59 formes, ce qui reste très peu.

Etant donné les formes que nous avons considérées (majoritairement des formes de catégories relevant de classes fermées), on peut aussi raisonnablement supposer que les 35 % restants du premier ensemble considéré seront majoritairement des formes de catégorie nominale, verbale, adjectivale et adverbiale (les polysyllabiques en *-ment*), soit des catégories relevant de classes ouvertes. Si l'on fait l'hypothèse que la distribution de la liason sur ces formes (restant à identifier) est identique à la distribution que nous avons identifiée (soit moins de 0,30% : 50 formes pour 17 775 contextes), les contextes non identifiés devraient se distribuer sur une trentaine de formes (dont 9 déjà identifiés), soit au final une vingtaine de formes supplémentaires. On arriverait alors à un ensemble comprenant environ 80 formes.

Ce qui se révèle par conséquent intéressant à travailler, est l'exploration, en détail, des conditions de réalisation de chacune des formes identifiées et ce, locuteur par locuteur. Mais se situer au niveau lexical implique de concevoir que la liason ne relève pas d'un processus phonologique général, mais pour une part importante de l'accident (du fait de la part importante de formes appartenant à des catégories relevant de classes fermées¹³⁸ mais aussi du fait de la variation) et par suite comporte un aspect mémoriel significatif. Nous posons qu'une telle approche, radicalement différente de ce qui a été envisagée jusqu'à maintenant, ouvre de nouvelles perspectives de recherche.

¹³⁸ En d'autres termes, nous y associons l'histoire « individuelle » de ces formes.

Références bibliographiques

- Adda-Decker, M. Boula de Mareüil, P. & Lamel, L. (1999). Pronunciation Variants in French: schwa & liaison. *ICPhS 99*, pp. 2239-2242.
- Ågren, J. (1973). *Enquête sur quelques liaisons facultatives dans le français de conversation radiophonique*. Uppsala : Acta Universitatis Upsaliensis.
- Anderson, J. R. (1983). *The Architecture of Cognition*. M.A., Cambridge : Harvard University Press.
- Armstrong, N. (2001). *Social and Stylistic Variation in Spoken French: a comparative approach*. Amsterdam : John Benjamins.
- Ashby, W. J. (1981). French Liaison as a Sociolinguistic Phenomenon. In William W. Cressey and Donna Jo Napoli (ed.), *Linguistic Symposium on Romance Languages 9*. Washington D.C. : Georgetown University Press, pp. 46-57.
- Auroux, S. (1998). *La raison, le langage et les normes*. Paris : PUF.
- Boltanski, J.-E. (1999). *Nouvelles Directions en Phonologie*. Paris : P.U.F.
- Bonami, O. & Boyé, G. (2003). La nature morphologique des allomorphies conditionnées : les formes des adjectifs en français. In B. Fradin *and al.* (ed.), *Les Unités morphologiques. Actes du 3^{ème} forum de morphologie*. Collection Silexicales : Université de Lille 3, pp. 39-48.
- Bonami, O., Boye, G. & Tseng, J. (2005). Sur la grammaire des consonnes latentes. *Langages*, n°158, pp. 89-100.
- Bonnard, H. (1979). *Synopsis de phonétique historique*. Paris : SEDES.
- Bourciez, E. & Bourciez, J. (1967). *Phonétique française, Etude historique*. Paris : Klincksieck.
- Bourciez, E. (1945). *Précis de phonétique française*. Paris : Klincksieck.
- Bourdieu, P. (1979). *La distinction. Critique sociale du jugement*. Collection Le sens commun. Paris : Minit.
- Brunot, F. (1966). *Histoire de la langue française*. Paris : Armand Colin. T. I, 1904 (rééd. 1966); t. II, 1906 (rééd. 1967) : 161-2; t. III-1, 1909 (rééd. 1966) : 81-94,

Références bibliographiques

- 262-71; t. IV-1, 1913 (rééd. 1966) : 25-49.
- Buben, V. (1935). *Influence de l'orthographe sur la prononciation du français moderne*. Bratislava : Faculté de lettres de l'université de Bratislava.
- Bybee, J. (1995). Regular morphology and the lexicon. *Languages and cognitive processes*, n°10, pp. 425-455.
- Bybee, J. (1998). A Functionalist Approach to Grammar and its Evolution. *Evolution of Communication*, n°2 (2), pp. 249-278.
- Bybee, J. (2000). The Phonology of the Lexicon: evidence from lexical diffusion. In M. Barlow and S. Kemmer (ed.), *Usage-Based Models of Language*. Stanford : CSLI Publications, pp. 65-85.
- Bybee, J. (2001a). *Phonology and Language Use*. Cambridge : Cambridge University Press.
- Bybee, J. (2001b). Frequency effects on French liaison. In J. Bybee and P. Hopper (ed.), *Frequency and the Emergence of Linguistic Structure*. Amsterdam : John Benjamins, pp. 337-359.
- Bybee, J. (2003). Mechanisms of change in grammaticization: the role of frequency. In B. D. Joseph and R. D. Janda (ed.), *The Handbook of Historical Linguistics*. Oxford : Blackwell Publishers, pp. 602-623.
- Bybee, J. (2005). La liaison : effets de fréquence et constructions. *Langages*, n°158, pp. 24-37.
- Bybee, J. (2007). *Frequency of Use and the Organization of Language*. Oxford : Oxford University Press.
- Cauchie, A. (1586). *Grammaire française*. Texte latin original. Traduction et notes de C. Demaizière. Paris : Champion, 2001.
- Chevrot, J.-P. (2001). *Variation phonétique, développement, orthographe*. Dossier non publié présenté en vue de l'Habilitation à diriger des recherches, 1 : Synthèse. Université Stendhal-Grenoble 3.
- Chevrot, J.-P. & Fayol, M. (2001). Acquisition of French liaison and related child errors. In M. Almgren and al. (ed.), *Research on child language acquisition*. Somerville, MA : Cascadilla Press, pp. 760-774.
- Chevrot, J.-P. & Malderez, I. (1999). L'effet Buben : De la linguistique diachronique à l'approche cognitive (et retour). *Langue Française*, n°158, pp. 38-52.
- Chevrot, J.-P., Chabanal, D. & Dugua, C. (2007). Pour un modèle de l'acquisition des

- liaisons basé sur l'usage : trois études de cas. *Journal of French Language Studies*, n°17, pp. 103-129.
- Chevrot, J.-P., Dugua, C. & Fayol, M. (2005). Liaison et formation des mots en français : un scénario développemental. *Langages*, n°158, pp. 38-52.
- Chevrot, J.-P., Fayol, M. & Laks, B. (ed.) (2005). La liaison : de la phonologie à la cognition. *Langages*, n°158, pp. 3-8.
- Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English*. New York : Harper and Row. Trad. française, P. Encrevé, *Principes de Phonologie générative*. Paris : Le Seuil, 1973.
- Clédat, L. (1917). *Manuel de phonétique et de morphologie*. Paris : Hachette.
- Clements, G. N. & Keyser, S. J. (1981). A three-Tired Theory of the Syllable. In C. F. C. Science (ed.), *Occasional Papers 19*. Cambridge : MIT Press.
- Clements, G. N. & Keyser, S. J. (1983). *CV Phonology, A Generative Theory of Syllable*. Cambridge Massachusetts : MIT Press.
- Cori, M. & David, S. (2008). Les corpus font-ils une nouvelle linguistique ? *Langages*, n°171, pp. 111-129.
- Coseriu, E. (1969). *Einführung in die strukturelle Linguistik*. Tübingen : Gunter Narr Verlag.
- Côté, M.-H. (2003). From representations to constraints in French liaison. *Communication, From representations to constraints*. Université de Toulouse-Le Mirail.
- Côté, M.-H. (2005). Le statut lexical des consonnes de liaison. *Langages*, n°158, pp. 66-78.
- Croft, W. A. (2001). *Radical Construction Grammar - Syntactic Theory in Typological Perspective*. Oxford : Oxford University Press.
- Dauzat, A. (1930). *Histoire de la langue française*. Paris : Payot.
- De Jong, D. (1988). *Sociolinguistic aspects of French Liaison*. Dordrecht : Foris.
- De Jong, D. (1994). La sociophonologie de la liaison orléanaise. In C. Lyche (ed.), *French Generative Phonology: Restrospective and Perspectives*. Salford : ESRI, pp. 95-131.
- Dejean de la Bâtie, B. (1993). *Word boundary ambiguity in spoken French*. Ph.D. dissertation : Monash University.

Références bibliographiques

- Delais-Roussarie, E. (2008). Corpus et données en phonologie post-lexicale. *Langages*, n°171, pp. 60-76.
- Delais-Roussarie, E., Durand, J., Lyche, C., Meqqori, A. & Tarrier, J.-M. (2003). Transcription des données : outils et conventions. *Bulletin PFC I*, pp. 21-34.
- Delattre, P. (1947). La liaison en français, tendances et classification. *The French Review*, n°22 (2), pp. 148-157.
- Delattre, P. (1951). *Principes de phonétique française à l'usage des étudiants anglo-américains*. Vermont : Middlebury College, 2^{ème} édition.
- Delattre, P. (1955). Les facteurs de la liaison facultative en français. *The French Review*, n°29 (1), pp. 42-49.
- Delattre, P. (1956). La fréquence de la liaison facultative en français. *The French Review*, n°30, pp. 48-54.
- Delattre, P. (1966). *Studies in French and comparative Linguistics*. The Hague : Mouton.
- Dell, F. (1970). *Les Règles phonologiques tardives et la Phonologie dérivationnelle du français*. Ph.D. dissertation : MIT.
- Dell, F. (1973). *Les Règles et les sons*. Paris : Hermann.
- Desrochers, R. (1994). Les liaisons dangereuses : le statut équivoque des erreurs de liaisons. *Linguisticae investigationes*, vol. 18, n°2, pp. 243-284.
- Dictionnaire de l'Académie française* (1762). Paris : Chez la Vve B. Brunet, 4^{ème} édition.
- Domergue, P. (1805). *Manuel des étrangers amateurs de la langue française*.
- Dugua, C. (2002). *Liaison et segmentation du lexique en français : vers un scénario développemental*. Université Stendhal.
- Dugua, C. (2006). *Liaison et segmentation lexicale et schémas syntaxiques entre 2 et 6 ans : un modèle développemental basé sur l'usage*. Ph.D. dissertation : Université de Grenoble III.
- Durand, J., Laks, B. & Lyche, C. (2001). *La Phonologie du français contemporain : usages, variétés et structure*. Document non publié.
- Durand, J., Laks, B. & Lyche, C. (2002). La phonologie du français contemporain : usages, variétés et structure. In C. Pusch and W. Raible (ed.), *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics-Corpora and Spoken Language*. Tübingen : Gunter Narr

Verlag, pp. 93-106.

- Durand, J. & Lyche, C. (2003). Le projet 'Phonologie du Français Contemporain' (PFC) et sa méthodologie. In E. Delais and J. Durand (ed.), *Corpus et variation en phonologie du français : méthodes et analyses*. Toulouse : Presses Universitaires du Mirail, pp. 212-276.
- Durand, J. & Lyche, C. (2004). Structure et variation dans quelques systèmes vocaliques du français : l'enquête phonologie du français contemporain. In A. Coveney, M.-A. Hintze and C. Sanders (ed.), *Variation et francophonie*. Paris : L'Harmattan, pp. 217-240.
- Durand, J. & Lyche, C. (à paraître). French liaison in the light of corpus data. *Journal of French Language Studies*, vol. 18/1.
- Durand, J. & TARRIER, J.-M. (2006). PFC, corpus et systèmes de transcription. *Cahiers de grammaire*, n°30, pp. 139-158.
- Encrevé, P. (1988). *La liaison avec et sans enchaînement. Phonologie tridimensionnelle et usages du français*. Paris : Seuil.
- Encrevé, P. (1997). L'ancien et le nouveau, quelques remarques sur la phonologie et son histoire. *Langages*, n°125, pp. 100-123.
- Estienne, H. (1582). *Hypomneses*. Paris.
- Eychenne, J. (2006). *Aspects de la phonologie du schwa dans le français contemporain : optimalité, visibilité prosodique, gradience*. Ph.D. dissertation : Université de Toulouse 2 Le Mirail.
- Fouché, P. (1935). La prononciation actuelle du français. In A. Dauzant (ed.), *Où en sont les études du français*. Paris : Bibliothèque du français moderne.
- Fouché, P. (1952). *Phonétique historique du français : Introduction*. Paris : Klincksieck.
- Fouché, P. (1959). *Traité de prononciation française*. Paris : Klincksieck.
- Fougeron, C., Goldman, J.-P. & Frauenfelder, U. (2001a). Liaison and schwa deletion in French: an effect of lexical frequency and competition? *Proceedings of Eurospeech 2001*. Aalborg : Danemark, pp. 639-642.
- Fougeron, C., Goldman J.P., Dart A., Guélat L. & Jeager C. (2001b). Influence de facteurs stylistiques, syntaxiques et lexicaux sur la réalisation de la liaison en français. *8ème Conférence Traitement Automatique des Langues Naturelles*. Tours, France.

Références bibliographiques

- Gadet, F. (1989). *Le français ordinaire*. Paris : Armand Colin Linguistique.
- Gadet, F. (2007). *La variation sociale en français*. Collection L'Essentiel français. Paris : Ophrys.
- Garmadi, J. (1981). *Sociolinguistique*. Paris : P.U.F.
- Goldberg, A. (1995). *Constructions : A Construction Grammar Approach to Argument Structure*. Chicago : CSLI.
- Goldsmith, J. A. (1976). *Autosegmental Phonology*. Ph.D. dissertation : MIT. New York : Garland Press, 1979.
- Gougenheim, G. (1935). *Eléments de phonologie française. Etude descriptive des sons du français au point de vue fonctionnel*. Paris : Les Belles Lettres.
- Gougenheim, G. (1938). *Le système grammatical de la langue française*. Paris : Bibliothèque du français moderne.
- Grammont, M. (1914). *Traité de prononciation française*. Paris : Delagrave.
- Grammont, M. (1933). *Traité de phonétique*. Paris : Delagrave.
- Green, J. & Hintze, M-A. (1988). Variation and change in French linking phenomena. In J. Green and W. Ayres-Bennet (ed.), *Variation and Change in French*. London : Routledge, pp. 61-88.
- Grevisse, M. (1969). *Le bon Usage*. Gembloux : Duculot.
- Hayes, B. (1999). Phonetically-driven phonology : the role of Optimality Theory and grounding. In M. Darnell and al. (ed.), *Functionalism and formalism in linguistics*, vol. 1, General papers. Amsterdam : Helmut Buske, pp. 243-286.
- Hindret, J. (1687). *L'Art de bien prononcer et de bien parler la langue française*. Réimpression Genève : Slatkine, 1973.
- Jong, T. de (2006). *La prononciation des consonnes dans le français de Paris aux 13ème et 14ème siècles*. Utrecht : LOT.
- Kager, R. (1999). *Optimality Theory*. Cambridge : Cambridge University Press.
- Kahn, D. (1976). *Syllable Based Generalizations in English Phonology*. Ph.D. dissertation : Massachusetts Institute of Technology.
- Kaye, J.D. & Lowenstam, J. (1984). De la Syllabicité. In F. Dell, D. Hirst and J-R. Vergnaud (ed.), *La forme sonore du langage*. Paris : Hermann, pp. 123-155.
- Kaye, J.D. & Morin, Y.C. (1976). Il n'y a pas de Règle de troncation, voyons...! In Dressler and Meid (ed.), *Proceedings of the 12th International Congress of*

Linguists. Germany : Innsbruck.

- Kilani-Schoch, M. (1983). Troncation ou insertion dans les liaisons françaises : quelques données aphasiques comme indices externes. *Folia linguistica*, n°17, pp. 445-461.
- Klausenburger, J. (1974). Rule inversion, opacity, conspiracies: French liaison and elision. *Lingua*, n°34, pp. 167-169.
- Klausenburger, J. (1976). French Linking Phenomena: A Natural Generative Analysis. Communication, *Summer Institute*. LSA : Oswego.
- Klausenburger, J. (1977). Deletion versus Epenthesis: intra-versus interparadigmatic arguments in linguistics. *Lingua*, n°42, pp. 153-160.
- Klausenburger, J. (1978a). French Linking Phenomena: A Natural Generative Analysis. *Language*, n°54, pp. 21-40.
- Klausenburger, J. (1978b). Liaison 1977: The Case for epenthesis. *Studies in French Linguistics*, n°1/2, pp. 1-20.
- Klausenburger, J. (1984). *French Liaison and linguistic theory*. Stuttgart : Franz Steiner Verlag Wiesbaden GMBH.
- Klausenburger, J. (2001). The Morphologization and Grammaticalization of French Liaison. In J. Herschensohn, E. Mallén and K. Zagana (ed.), *Features and Interfaces in Romance*.
- Kurylowicz, J. (1948). Contribution à la théorie de la syllabe. *Bulletin de la société polonaise de linguistique*, VIII, pp. 80-114.
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia : U. of Pennsylvania Press. Trad. française : *Sociolinguistique*. Paris : Editions de Minuit, 1972.
- Labov, W. (1973). *Language in the Inner City*. Trad. française : *Le Parler ordinaire*. Paris : Editions de Minuit, 1979, 2 vol.
- Labov, W. (2002). Qu'est-ce qu'un fait linguistique? In M. Santacrose (ed.), *Faits de langue-Faits de discours. Données, processus et modèles. Qu'est-ce qu'un fait linguistique ?* n°1. Paris : L'Harmattan, pp. 11-117.
- Labrone, L. (2005). Autour de la syllabe : les constituants prosodiques mineurs en phonologie. In. N. Nguyen, S. Wauquier and J. Durand (ed.), *Phonétique et phonologie, approches contemporaines*. Paris : Hermès, pp. 95-116.
- Laks, B. (1977). Contribution empirique à l'analyse socio-différentielle de la chute des /r/ dans les groupes consonantiques finals. *Langue française*, n°34, pp. 109-125.

Références bibliographiques

- Laks, B. (1980). Différenciation linguistique et différenciation sociale : quelques problèmes de sociolinguistique française. Ph.D. dissertation : Paris VIII.
- Laks, B. (2005a). Phonologie et construction syntaxique : la liaison, un test de figement et de cohésion. *Linx*, n°53, pp. 155-171.
- Laks, B. (2005b). La liaison et l'illusion. *Langages*, n°158, pp. 101-126.
- Langacker, R. (1991). *Foundations of Cognitive Grammar*. Stanford : Stanford University Press, 2ème édition.
- Langlard, H. (1928). *La Liaison dans le français*. Paris : Champion.
- Laveaux, C. (1847). *Dictionnaire raisonné des difficultés grammaticales et littéraires de la langue française*. Paris : A. Ledentu fils, 3^{ème} édition.
- Littré, P.-E. (1866). *Dictionnaire de la langue française*. Paris : Encyclopédia Britannica, 1974.
- Lodge, R. A. (1997). *Le français : histoire d'un dialecte devenu langue*. Paris : Fayard.
- Long, M. (1978). Comments on Klausenburger's Paper. *Studies in French Linguistics*, vol. ½, pp. 21-26.
- Lucci, V. (1983). Prosodie, phonologie et variation en français contemporain. *Langue française*, n°60, pp. 73-84.
- Lyche, C. (2004). Le corpus pfc (phonologie du français contemporain) pour un renouvellement des données phonologiques. *Cahiers du LIDL*, n°31.
- Lyche, C. (2005). Liaison in PFC, what can we learn from the data? *7e colloque du GDR 1954 Phonologie*. 2-4 June, Aix-en-Provence.
- Malécot, A. (1975). French Liaison as a Fonction of Grammatical, Phonetic and Paralinguistic Variables. *Phonetica*, n°32, pp. 161-179.
- Mallet, G.-M. (2003). Quelques observations sur la liaison à Treize-Vents (Vendée). *La Tribune Internationale des Langues Vivantes*, n° 33, pp. 145-150.
- Martinet, A. (1945). *La prononciation du français contemporain*. Genève : Droz, 2^{ème} édition.
- Martinet, A. (1964). *L'économie des changements phonétiques : Traité de phonologie diachronique*. Berne : Francke.
- Martinon, P. (1913). *Comment on prononce le français. Traité complet de prononciation pratique avec les noms propres et les noms étrangers*. Paris : Larousse.

- McCarthy, J. (2002). Comparative Markedness (long version). *Papers in Optimality Theory*, II, pp. 171-246.
- McCarthy, J. & Prince, A. (1993a). Generalized Alignment. *Yearbook of Morphology*. Dordrecht : Kluwer, pp. 79-153.
- McCarthy, J. & Prince, A. (1993b). *Prosodic Morphology: Constraint Interaction and Satisfaction*. Rapport technique RuCSS-TR-3/. New Brunswick : ROA.
- McCarthy, J. & Prince, A. (1995). Faithfulness and Identity in Prosodic Morphology. In J. Beckman, L. W. Dickey and S. Urbanczyk (ed.), *University of Massachusetts Occasional Papers in Linguistics*. Amherst : GLSA, pp. 249-384.
- Meigret, L. (1542). *Traité touchant le commun usage de l'écriture française*. Lyon : Republications Slatkine Genève, 1972.
- Milner, J.-C. (1967). French Truncation Rule. *Quarterly Progress Report of the Research Laboratory of Electronics*, MIT 86, pp. 273-283.
- Milner, J.-C. (1989). *Introduction à une science du langage*. Paris : Le Seuil.
- Milroy, J. (1980). *Language and Social Networks*. Oxford : Blackwell.
- Morel, E. (1994). Le traitement de la liaison chez l'enfant : études expérimentales. *Travaux neuchâtelois de linguistique*, n°21, pp. 85-95.
- Morin, Y.-C. (1979a). La morphophonologie des pronoms clitiques en français populaire. *Cahier de linguistique*, n°9, pp. 1-36.
- Morin, Y.-C. (1979b). More remarks on French clitic order. *Linguistic analysis*, n°5, pp. 293-312.
- Morin, Y.-C. (1981). Où sont passés les s finals de l'ancien français? In D. Sankoff and H. J. Cedergren (ed.), *Variation omnibus*. Edmonton : Linguistic research, pp. 35-47.
- Morin, Y.-C. (1987). French data and phonological theory. *Linguistics*, n°25, pp. 815-843.
- Morin, Y.C. (1994). Phonological interpretation of historical lengthening. In W. U. Dressler, M. Prinzhorn and J. Rennison (ed.), *Phonologica 1992, Proceedings of the 7th International phonology meeting*. Turin : Rosenberg & Sellier, pp. 135-155.
- Morin, Y.-C. (2000). Le français de référence et les normes de prononciation. *Cahiers de l'Institut de Linguistique de Louvain*, n°26 (1), pp. 91-135.
- Morin, Y.-C. (2003). Remarks on prenominal liaison consonants in French. In S. Ploch (ed.), *Living on the edge. 28 papers in honour of Jonathan Kaye*. Berlin : Mouton

Références bibliographiques

- de Gruyter, pp. 385-400.
- Morin, Y-C. (2005). La liaison relève-t-elle d'une tendance à éviter les hiatus. Réflexions sur son évolution historique. *Langages*, n°158, pp. 8-23.
- Morin, Y-C. & Kaye J. (1982). The Syntactic Bases for French Liaison. *Journal of Linguistics*, n°18, pp. 291-330.
- Mounin, G. (1974). *Dictionnaire de la linguistique*. Paris : Presses Universitaires de France.
- Nyrop, K. (1934). *Manuel phonétique du français parlé*. Copenhague : Nordiske.
- Pagliano, C. (2003). *L'épenthèse consonantique en français. Ce que la syntaxe, la sémantique et la morphologie peuvent faire à la phonologie : parles-en de ta numérotation automatique*. Ph.D. dissertation : Université de Nice-Sophia Antipolis.
- Pagliano, C. & Laks, B. (2005). Problématique de la liaison dans l'analyse d'un corpus de français oral actuel. *Français fondamental, corpus oraux, contenus d'enseignement. 50 ans de travaux et d'enjeux*, 8-10 décembre 2005. Lyon : École Normale Supérieure Lettres et Science Humaines.
- Paradis, C. & El Fenne, F. (1995). French verbal inflection revisited : Constraints, repairs and floating consonants. *Lingua*, n°95, pp. 169-204.
- Passy, P. (1892). *Les sons du français. Leur formation, leur combinaison, leur représentation*. Paris : Firmin-Didot.
- Perlmutter, D. (1998). Interfaces : explanation of allomorphy and the architecture of grammars. In G. Lapointe and al. (ed.), *Morphology and its relation to phonology and syntax*. Stanford : CSLI, pp. 307-338.
- Pierrehumbert, J. (1999). Formalizing functionalism. In M. Darnell and al. (ed.), *Functionalism and formalism in linguistics*, vol. 1, General papers. Amsterdam : Helmut Buske, pp. 287-304.
- Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven and N. Warner (ed.), *Laboratory Phonology VII*. Berlin : Mouton de Gruyter, pp. 101-140.
- Pierrehumbert, J., Beckmann, M.E. & Ladd, D.R. (2000). Conceptual foundations of phonology as a laboratory science. In N. Burton-Roberts, P. Carr and G. Docherty (ed.), *Phonological knowledge, Conceptual and Empirical Issues*. Oxford : University Press, pp. 273-305.
- Pike, K. & Pike, L. (1947). Immediate Constituents of Mazateco Syllables. *International Journal of American Linguistics*, n°13, pp. 78-91.

- Post, B. (2000). Pitch Accents, Liaison and the Phonological Phrase in French. *Probus*, n°12, pp. 127-164.
- Prince, A. & Smolensky, P. (1993). Optimality Theory : Constraint Interaction in Generative Grammar. ROA.
- Prince, A. & Smolensky, P. (1997). Optimality: From neural networks to universal grammar. *Science* 275, pp. 1604-1610.
- Pulgram, E. (1970). *Syllable, Word, Nexus, Cursus*. The Hague : Mouton.
- Ranson, D. (2008). La liaison variable dans un corpus du français méridional : L'importance relative de la fonction grammaticale. *Congrès mondial de linguistique française*, pp. 1689-1683.
- Saussure F. de (1916). *Cours de linguistique générale*. Paris: Payot.
- Schane, S. A. (1965). *The Phonological and Morphological Structure of French*, Ph.D. dissertation : MIT Press.
- Schane, S. A. (1967). La Liaison et l'élision en français. *Langages*, n°8, pp. 37-59.
- Schane, S. A. (1968). *French Phonology and Morphology*. Cambridge, Massachusetts : M.I.T. Press.
- Schane, S. A. (1974). There is no French Truncation Rule. In R. Campbell, M. G. Goldin and M. Clayton Wang (ed.), *Linguistic studies in Romance Languages*. Washington D.C. : Georgetown University Press, pp. 89-100.
- Selkirk, E. (1972). *The Phrase Phonology of English and French*. Ph.D. dissertation : MIT. New York : Garland, 1980.
- Selkirk, E. (1974). French Liaison and X notation. *Linguistic Inquiry*, n°5, pp. 573-590.
- Selkirk, E. (1981). Epenthesis and Degenerate Syllables in Cairene Arabic. *MIT Working Papers in Linguistics*, n°3, pp. 209-232.
- Selkirk, E. (1984). On the Major Class Features and Syllable Theory. In M. Aronoff and R. T. Oehrle (ed.), *Language Sound Structure*. Cambridge, Massachusetts : MIT Press, pp. 107-136.
- Sobotta, E. (2006). *Phonologie et migration – Aveyronnais et Guadeloupéens à Paris*. Ph.D. dissertation : Paris X-Nanterre et Ludwig-Maximilians-Universität.
- Spinelli, E., Mc Queen, J.M. & Cutler, A. (2003). Processing resyllabified words in French. *Journal of memory and language*, n°48, pp. 233-254.
- Steriade, D. (1999). Lexical Conservatism in French Adjectival Liaison. In B. Bullock,

Références bibliographiques

- M. Authier and L. Reed (ed.), *Formal Perspectives in Romance Linguistics*. Amsterdam : John Benjamins, pp. 243-270.
- Ternes, E. (1977). Konsonantische Anlautveränderungen in den keltischen und romanischen Sprachen. *Romanistisches Jahrbuch*, n°28, pp. 19-53.
- Thimonnier, R. (1970). *Code orthographique et grammatical*. Paris : Hatier.
- Thurot, C. (1881-1883). *De la prononciation française depuis le commencement du XVIe siècle, d'après les témoignages des grammairiens*, 2 vol. Paris : Imprimerie Nationale. Réimpression, Genève : Slatkine, 1973.
- Tranel, B. (1974). *The Phonology of Nasal Vowels in Modern French*. Ph.D Dissertation : University of California.
- Tranel, B. (1976). A note of final consonant deletion in modern French. *Lingua*, n°39, pp. 53-68.
- Tranel, B. (1978). The Status of Nasal Vowels in Modern French. *Studies in French Linguistics*, vol. 1, pp. 27-70.
- Tranel, B. (1980). The [al] ~ [o] alternation in Modern French. In M. Aronoff and M.-L. Kean (ed.), *Juncture*. Saratoga : Anna Libri.
- Tranel, B. (1981). *Concreteness in Generative Phonology: evidence from French*. Berkeley : University of California Press.
- Tranel, B. (1986). French liaison and extrasyllabicity. In O. Jaeggli and C. Silva-Corvalan (ed.), *Studies in Romance linguistics*. Dordrecht : Foris, pp. 283-395.
- Tranel, B. (1990). On suppletion and French liaison. *Probus*, n°2, pp. 169-208.
- Tranel, B. (1995a). Current Issues in French Phonology: Liaison and Position Theories. In J. A. Goldsmith (ed.), *The Handbook of Phonological Theory*. Oxford UK and Cambridge MA : Blackwell, pp. 798-816.
- Tranel, B. (1995b). French final consonants and nonlinear phonology. *Lingua*, n°95, pp. 131-167.
- Tranel, B. (1998). Questioning Generalized Suppletion, communication. *Fourth Southwestern Workshop on Optimality Theory*. University of Arizona, Tucson.
- Tranel, B. (1999). Suppletion and OT: On the issue of the syntax/phonology interaction. In E. Curtis and al. (ed.), *Proceedings of the Sixteenth West Coast Conference on Formal Linguistics*. Stanford, CA : CSLI Publications, pp. 415-429.
- Tranel, B. (2000). Aspects de la phonologie du français et la théorie de l'optimalité. *Langue française*, n°126, pp. 39-72.

- Tribouillard, J. (1998). De l'hiatus en en poésie. *Défense de la langue française*, n°188, avril-mai-juin 1998, pp. 17-19.
- Trudeau, D. (1992). *Les Inventeurs du bon usage*. Paris : Minuit.
- Turcsan, G. (2005). *Le mot phonologique en français du Midi. Domaines, contraintes, opacité*. Ph.D. dissertation : Université Toulouse Le Mirail.
- Väänänen, V. (1962). *Introduction au latin vulgaire*. Paris : Klincksieck.
- Vaugelas, C. Fabre. de (1647). *Remarques sur la langue française*. Paris : Droz.
- Vergnaud, J.-R. (1982). Communication orale, sans titre, en anglais, communication, *GLOW Meeting* : Paris.
- Wauquier-Gravelines, (2001). Du réalisme des formalisations phonologiques contemporaines : que nous apprennent les données d'acquisition ? In J.P. Angoujard and S. Wauquier-Gravelines (ed.), *Phonologie : champs et perspectives*. Presses universitaires de l'Ecole Normale Supérieure de Fontenay-St-Cloud.
- Wauquier-Gravelines, S. (2005). *Statut des représentations phonologiques en acquisition, traitement de la parole continue et dysphasie développemental*. Paris : EHESS.
- Wauquier-Gravelines, S. & Braud, V. (2005). Proto-déterminant et acquisition de la liaison obligatoire en français. *Langage*, n° 158. Paris : Larousse.
- Wetzels, W. L. (1987). The Timing of Latent Consonants in Modern French. In C. Neidle and R. A. Nuñez Cedeño, *Studies in Romance Languages*. Dordrecht, Providence : Foris Publications, pp. 283-317.
- Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort*. Addison-Wesley.

Liste des annexes

Annexe 1 Texte

Le Premier Ministre ira-t-il à Beaulieu ?

Le village de Beaulieu est en grand émoi. Le Premier Ministre a en effet décidé de faire étape dans cette commune au cours de sa tournée de la région en fin d'année. Jusqu'ici les seuls titres de gloire de Beaulieu étaient son vin blanc sec, ses chemises en soie, un champion local de course à pied (Louis Garret), quatrième aux jeux olympiques de Berlin en 1936, et plus récemment, son usine de pâtes italiennes. Qu'est-ce qui a donc valu à Beaulieu ce grand honneur ? Le hasard, tout bêtement, car le Premier Ministre, lassé des circuits habituels qui tournaient toujours autour des mêmes villes veut découvrir ce qu'il appelle « la campagne profonde ».

Le maire de Beaulieu – Marc Blanc – est en revanche très inquiet. La cote du Premier Ministre ne cesse de baisser depuis les élections. Comment, en plus, éviter les manifestations qui ont eu tendance à se multiplier lors des visites officielles ? La côte escarpée du Mont Saint-Pierre qui mène au village connaît des barrages chaque fois que les opposants de tous les bords manifestent leur colère. D'un autre côté, à chaque voyage du Premier Ministre ; le gouvernement prend contact avec la préfecture la plus proche et s'assure que tout est fait pour le protéger. Or, un gros détachement de police, comme on en a vu à Jonquière ; et des vérifications d'identité risquent de provoquer une explosion. Un jeune membre de l'opposition aurait déclaré : « Dans le coin, on est jaloux de notre liberté. S'il faut montrer patte blanche pour circuler, nous ne répondrons pas de la réaction des gens du pays. Nous avons le soutien du village entier. » De plus, quelques articles parus dans La Dépêche du Centre, L'Express, Ouest Liberté et Le Nouvel Observateur indiqueraient que des activistes des communes voisines préparent une journée chaude au Premier Ministre. Quelques fanatiques auraient même entamé un jeûne prolongé dans l'église de Saint Martinville.

Le sympathique maire de Beaulieu ne sait plus à quel saint se vouer. Il a le sentiment de se trouver dans une impasse stupide. Il s'est, en désespoir de cause, décidé à écrire au Premier Ministre pour vérifier si son village était vraiment une étape nécessaire dans la tournée prévue. Beaulieu préfère être inconnue et tranquille plutôt que de se trouver au centre d'une bataille politique dont, par la télévision, seraient témoins des millions d'électeurs.

Annexe 2 Liste de mots

- | | | |
|--------------------|--------------------|------------|
| 1. roc | 44. reliure | 87. épais |
| 2. rat | 45. piqué | 88. épée |
| 3. jeune | 46. malle | 89. jeune |
| 4. mal | 47. gnôle | 90. jeûne |
| 5. ras | 48. bouleverser | 91. beauté |
| 6. fou à lier | 49. million | 92. botté |
| 7. des jeunets | 50. explosion | 93. brun |
| 8. intact | 51. influence | 94. brin |
| 9. nous prendrions | 52. mâle | |
| 10. fêtard | 53. ex-mari | |
| 11. nièce | 54. pomme | |
| 12. pâte | 55. étrier | |
| 13. piquet | 56. chemise | |
| 14. épée | 57. brin | |
| 15. compagnie | 58. lierre | |
| 16. fête | 59. blanc | |
| 17. islamique | 60. petit | |
| 18. agneau | 61. jeûne | |
| 19. pêcheur | 62. rhinocéros | |
| 20. médecin | 63. miette | |
| 21. paume | 64. slip | |
| 22. infect | 65. compagne | |
| 23. dégeler | 66. peuple | |
| 24. bêttement | 67. rauque | |
| 25. épier | 68. cinquième | |
| 26. millionnaire | 69. nier | |
| 27. brun | 70. extraordinaire | |
| 28. scier | 71. meurtre | |
| 29. fêter | 72. vous prendriez | |
| 30. mouette | 73. botté | |
| 31. déjeuner | 74. patte | |
| 32. ex-femme | 75. étriller | |
| 33. liège | 76. faites | |
| 34. baignoire | 77. feutre | |
| 35. pêcheur | 78. quatrième | |
| 36. socialisme | 79. muette | |
| 37. relier | 80. piquais | |
| 38. aspect | 81. trouer | |
| 39. niais | 82. piquer | |
| 40. épais | 83. creuse | |
| 41. des genêts | 84. beauté | |
| 42. blond | 85. patte | |
| 43. creux | 86. pâte | |

Annexe 3 Fiche signalétique

NOM, prénom(s) :
Date de naissance :
Lieu de naissance :
Domiciles successifs (en nombre d'années)
.....
Domicile actuel :
.....
Professions successives :
.....
Profession actuelle :
Situation familiale (marié, célibataire, etc.) :
Enfants, âge, scolarisation :
.....
Études (préciser jusqu'à quel âge et quel type d'études) :
.....
Langues parlées :
.....

Père de l'informateur, année de naissance :
Lieu d'origine :
Profession :
Etudes :
Langues parlées (étrangères ou régionales) :

Mère de l'informateur, année de naissance :
Lieu d'origine :
Profession :
Etudes :
Langues parlées (étrangères ou régionales) :

Époux, épouse, autre :
Lieu d'origine :
Profession :
Etudes :
Langues :
Personnes ayant joué un rôle important au moment de l'apprentissage du français par l'enquêté (grands-parents, nourrice...) :
.....

Type de logement de l'enquêté (maison, appartement...) :
Intégration dans le quartier, relations de voisinage :
Activités culturelles, loisirs, voyages :
Autres informations :
.....
.....

Nom de l'enquêteur (entretien guidé) :
Nom de l'enquêteur (entretien libre) :
Date de l'enregistrement, durée :

Liste des annexes

Lieu de l'enregistrement :
Localité :
Lien enquêteur/enquêté :
Professionnel :
Amical :
Familial :
Autre :
Observations sur l'enregistrement :
Informel : qui étaient les locuteurs présents ?
Ordre des situations dans l'enregistrement (ex. guidé, liste de mots, texte, libre) :
.....
Principaux thèmes abordés :
.....
Qualité de l'enregistrement :
Autres observations : (interventions d'autres personnes, interruptions téléphoniques notables...)
.....
.....

Enregistrement de groupe

Établir une fiche signalétique pour chaque participant et préciser les relations à l'intérieur du Groupe.

Annexe 4 Consentement de participation

NOM :

Prénom :

M / Mme X vous propose de participer à une enquête dans le cadre du projet universitaire : « La phonologie du français contemporain : usages, variétés et structures ». Ce projet s'appuie sur des enregistrements effectués sur des échantillons de francophones de divers pays du monde.

Ces enregistrements sont faits dans un but de recherche scientifique. Les résultats (y compris les enregistrements) peuvent en être publiés aussi bien dans des revues scientifiques que dans des ouvrages commercialisés (par exemple, manuels pédagogiques). Ils pourront éventuellement être intégrés à des systèmes de production et de perception de la parole. Dans tous les cas de publication, les informations vous concernant seront exploitées de manière anonyme. Vous pourrez par ailleurs avoir accès à toute publication éventuelle si vous en faites la demande.

Votre consentement ne décharge pas les enquêteurs de leur responsabilité. Vous conservez tous vos droits garantis par la loi.

Si vous acceptez de faire partie de l'échantillon d'études dans les conditions énoncées ci-dessus, veuillez signer ce document.

Signature

Annexe 5 Répartition des locuteurs par secteur et par sexe

Secteur	Enquêtes	Femmes	Hommes
Midi	Aix-Marseille 13b	3	5
	Biarritz 64a	7	5
	Douzens 11a	5	5
	Lacaune 81a	8	5
	Rodez 12a	5	4
	Toulouse 31a	7	7
Total		35	31
Nord	Aveyronnais à Paris 75x	7	5
	Brécey 50a	6	5
	Brunoy 91a	5	5
	Dijon 21a	5	3
	Grenoble 38a	9	3
	Lyon 69a	6	5
	Nantes 44a	6	5
	Ogéville 54b	7	5
	Paris Centre ville 75c	6	6
	Puteaux-Courbevoie 92a	3	3
	Roanne 42a	4	5
	Vendée (85a)	3	5
Total		67	56
International	Abidjan cia	7	7
	Gembloux bga	6	6
	Ile de la Réunion 974	5	4
	Liège bla	6	6
	Nyon sva	5	7
	Tournai bta	6	6
Total		35	36

Annexe 6 Répartition des locuteurs par tranche d'âge

Enquêtes	10 - 30 ans	30 - 45 ans	45 - 60 ans	> 60 ans	Total
Abidjan cia	6	3	5	0	14
Aix-Marseille 13b	2	0	4	2	8
Aveyron. à Paris 75x	6	2	1	3	12
Biarritz 64a	2	6	0	4	12
Brécey 50a	3	4	2	2	11
Brunoy 91a	2	0	3	5	10
Dijon 21a	6	1	0	1	8
Douzens 11a	3	0	3	4	10
Gembloux bga	4	1	4	3	12
Grenoble 38a	6	1	2	3	12
Ile de la Réunion 974	4	1	2	2	9
Lacaune 81a	2	2	2	7	13
Liège bla	4	3	1	4	12
Lyon 69a	5	2	3	1	11
Nantes 44a	3	1	3	4	11
Nyon sva	2	3	5	2	12
Ogéville 54b	3	1	3	5	12
Paris 75c	4	3	2	3	12
Puteaux 92a	4	0	2	0	6
Roanne 42a	0	0	2	7	9
Rodez 12a	4	0	2	3	9
Toulouse 31a	12	2	0	0	14
Tournai bta	5	0	4	3	12
Vendée 85a	2	2	2	2	8
Total	94	38	57	70	259
%	36,29 %	14,67 %	22,00 %	27,02 %	100 %

Annexe 7 Occurrences de liaisons codées « incertaines »

	Locuteurs	Liaisons incertaines		
C O N V E R S A T I O N S	50apb1	de devises	23	entre l'euro et le dollar
	13bpa1	Des	13	expressions provençales.
	75xrg1	et puis les	13	autres bricoles
	abahs1	et quand	13	vous êtes
	81aar3	et s'assure que tout	13	est, fait pour, le protéger.
	81aac1	Eux	13	enfin oui que,
	12atp1	il y en	13	a/ avait pas beaucoup
	12ajp1	j'en	13	ai connu oui.
	81ajc1	les organismes	23	agricoles hein.
	91acs2	Lui, ça allait	13	être dix- huit.
	12aja1	maintenant il y en	13	a peut-être trente
	12ags1	nés là-bas en	13	Algérie.
	64ama4	nous re/ nous	13	é/ nous étions
	64ajc1	pour moi c'est	13	une évolution
	50aid1	que des	13	activi/ activismes
	64ama4	qui s'en vont	13	au (XXX)
	svarv1	qui sont des	13	émissions
svajb1	sans qu'elle soit	13	obligée de travailler.	
91ael1	tantôt dans	13	un endroit tantôt	
T E X T E	50arm1	de se trouver	23	au centre d'une bataille politique
	12acr1	de se trouver dans	13	une impasse stupide
	81aar2	Il s'est	13	en désespoir de cause
	abayt1	Il s'est	13	en désespoir de cause
	svarb2	Il s'est	13	en désespoir de cause
	svayb1	Il s'est	13	il s'est, en désespoir de cause
	92aaf3	des circuits	23	habituels
	92aaf1	des circuits	23	habituels
	81amb1	ses chemises	23	en soie
	31aec1	ses chemises	23	en soie
	abafb1	lors des visites	23	officielles

Annexe 8 Contextes de liaison identifiés dans le texte PFC

Catégoriques	Variables	Erratiques
<i>grand_émoi</i>	<i>est_en</i>	<i>Berlin_en</i>
<i>en_effet</i>	<i>chemises_en soie</i>	<i>comment_en plus</i>
<i>jeux_olympiques</i>	<i>pâtes_italiennes</i>	<i>le coin_on</i>
<i>son_usine</i>	<i>circuits_habituels</i>	
<i>grand_honneur</i>	<i>toujours_autour</i>	
<i>très_inquiet</i>	<i>est_en revanche</i>	
<i>les_élections</i>	<i>ont_eu</i>	
<i>les_opposants</i>	<i>visites_officielles</i>	
<i>un_autre côté</i>	<i>provoquer_une</i>	
<i>tout_est</i>	<i>préparent_une j.</i>	
<i>on_en</i>	<i>fanatiques_auraient</i>	
<i>en_a vu</i>	<i>plus_à quels</i>	
<i>on_est</i>	<i>s'est_en</i>	
<i>nous_avons</i>	<i>trouver_au</i>	
<i>quelques_articles</i>	<i>vraiment_une étape</i>	
<i>des_activistes</i>		
<i>dans_une</i>		

Annexe 9 Liste des formes identifiées

Les cinquante formes identifiées sont triées par ordre alphabétique.

	Forme	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
1	<i>allait</i>	45	2	43	4,4 %
2	<i>ans</i>	130	0	130	0 %
3	<i>assez</i>	41	2	39	4,88 %
4	<i>avaient</i>	408	6	402	1,5 %
5	<i>avait</i>	374	10	364	2,7 %
6	<i>avons</i>	24	5	19	20,83 %
7	<i>bien</i>	185	79	106	42,7 %
8	<i>c'est</i>	1 470	413	1 057	28,1 %
9	<i>chez</i>	64	53	11	82,81 %
10	<i>dans</i>	398	378	20	94,97 %
11	<i>des</i>	796	769	27	96,61 %
12	<i>deux</i>	334	334	0	100 %
13	<i>en</i>	959	939	20	97,9 %
14	<i>est</i>	636	279	357	43,9 %
15	<i>étaient</i>	71	8	63	11,3 %
16	<i>était</i>	343	29	314	8,5 %
17	<i>êtes</i>	16	5	11	31,25 %
18	<i>fallait</i>	35	0	35	0 %
19	<i>faut</i>	127	0	127	0 %
20	<i>grand</i>	9	5	4	55,6 %
21	<i>grandes</i>	5	4	1	80 %
22	<i>grands</i>	7	5	2	71,43 %
23	<i>ils</i>	864	853	11	98,73 %
24	<i>les</i>	1 047	1 030	17	98,38 %
25	<i>mais</i>	791	4	787	0,51 %
26	<i>mes</i>	148	148	0	100 %
27	<i>mon</i>	124	124	0	100 %
28	<i>nous</i>	427	320	107	74,94 %
29	<i>on</i>	2 489	2 489	0	100 %
30	<i>ont</i>	178	17	161	9,6 %
31	<i>pas</i>	880	12	868	1,36 %
32	<i>petit</i>	48	36	12	75 %
33	<i>petites</i>	13	12	1	92,31 %
34	<i>petits</i>	22	9	13	40,91 %
35	<i>peut</i>	94	29	65	30,9 %
36	<i>plus</i>	248	159	89	64,11 %
37	<i>puis</i>	639	1	638	0,16 %
38	<i>quand</i>	639	498	141	77,9 %
39	<i>ses</i>	62	62	0	100 %

Liste des annexes

40	<i>sommes</i>	27	12	15	44,44 %
41	<i>son</i>	58	58	0	100 %
42	<i>sont</i>	208	40	168	19,2 %
43	<i>suis</i>	430	58	372	13,49 %
44	<i>toujours</i>	184	0	184	0 %
45	<i>tout</i>	287	238	49	82,9 %
46	<i>très</i>	145	140	5	96,55 %
47	<i>trois</i>	211	182	29	86,26 %
48	<i>un</i>	551	540	11	98 %
49	<i>vais</i>	51	2	49	3,92 %
50	<i>vous</i>	433	425	8	98,15 %
	Total	17 775	10 823	6 952	

Annexe 10 Contextes de liaisons en /z/ identifiés

Les vingt-huit formes contenant la CL /z/ sont classées par ordre alphabétique.

	Forme	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
1	<i>ans</i>	130	0	130	0 %
2	<i>assez</i>	41	2	39	4,88 %
3	<i>avons</i>	24	5	19	20,83 %
4	<i>chez</i>	64	53	11	82,81 %
5	<i>dans</i>	398	378	20	94,97 %
6	<i>des</i>	796	769	27	96,61 %
7	<i>deux</i>	334	334	0	100 %
8	<i>êtes</i>	16	5	11	31,25 %
9	<i>grandes</i>	5	4	1	80 %
10	<i>grands</i>	7	5	2	71,43 %
11	<i>ils</i>	864	853	11	98,73 %
12	<i>les</i>	1 047	1 030	17	98,38 %
13	<i>mais</i>	791	4	787	0,51 %
14	<i>mes</i>	148	148	0	100 %
15	<i>nous</i>	427	320	107	74,94 %
16	<i>pas</i>	880	12	868	1,36 %
17	<i>petites</i>	13	12	1	92,31 %
18	<i>petits</i>	22	9	13	40,91 %
19	<i>plus</i>	248	159	89	64,11 %
20	<i>puis</i>	639	1	638	0,16 %
21	<i>ses</i>	62	62	0	100 %
22	<i>sommes</i>	27	12	15	44,44 %
23	<i>suis</i>	430	58	372	13,49 %
24	<i>toujours</i>	184	0	184	0 %
25	<i>très</i>	145	140	5	96,55 %
26	<i>trois</i>	211	182	29	86,26 %
27	<i>vais</i>	51	2	49	3,92 %
28	<i>vous</i>	433	425	8	98,15 %
	Total	8 437	4 984	3 453	59,07 %

Annexe 11 Contextes de liaisons en /n/ identifiés

Les 6 formes contenant la CL /n/ sont classées par ordre alphabétique.

	Forme	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
1	<i>bien</i>	185	79	106	42,7 %
2	<i>en</i>	959	939	20	97,9 %
3	<i>mon</i>	124	124	0	100 %
4	<i>on</i>	2 489	2 489	0	100 %
5	<i>son</i>	58	58	0	100 %
6	<i>un</i>	551	540	11	98 %
	Total	8 437	4 984	3 453	59,07 %

Annexe 12 Contextes de liaisons en /t/ identifiés

Les seize formes contenant une CL en /t/ sont classées par ordre alphabétique.

	Forme	Nombre en contexte de L	Nombre de LR	Nombre de LNR	% de LR
1	<i>allait</i>	45	2	43	4,4 %
2	<i>avaient</i>	408	6	402	1,5 %
3	<i>avait</i>	374	10	364	2,7 %
4	<i>c'est</i>	1 470	413	1 057	28,1 %
5	<i>est</i>	636	279	357	43,9 %
6	<i>étaient</i>	71	8	63	11,3 %
7	<i>était</i>	343	29	314	8,5 %
8	<i>fallait</i>	35	0	35	0 %
9	<i>faut</i>	127	0	127	0 %
10	<i>grand</i>	9	5	4	55,6 %
11	<i>ont</i>	178	17	161	9,6 %
12	<i>petit</i>	48	36	12	75 %
13	<i>peut</i>	94	29	65	30,9 %
14	<i>quand</i>	639	498	141	77,9 %
15	<i>sont</i>	208	40	168	19,2 %
16	<i>tout</i>	287	238	49	82,9 %
	Total	4 972	1 610	3 362	32,4 %

Annexe 13 Ensemble des contextes de liaisons erratiques

	M1	M2
Liaisons erratiques	nom singulier	
	nom propre	
	pronom (pas personnel) nasal	
	pronom personnel postposé nasal 3ème personne	- pronom personnel sujets (<i>on, ils, elles</i>) - pronom complément (<i>en, les</i>)
	adverbe interrogatif	
	conjonction polysyllabique	
	conjonction <i>et</i>	
		<i>h aspiré</i>
		<i>oui, non</i>
	- adjectif singulier - pronom nasal	invariable
		- <i>un, huit, onze</i> - les dérivés : <i>unième, huitième, onzième</i>
	nom composé au pluriel	
	groupe figé sans liaison	

Annexe 14 Contextes de liaison où les deux catégories (M1 et M2) sont spécifiées

	M1	M2
Liaisons erratiques	pronom personnel postposé nasal 3ème personne	- pronom personnel sujets (<i>on, ils, elles</i>) - pronom complément (<i>en, les</i>)
	- adjectif singulier - pronom nasal	invariable
	nom pluriel	verbe
Liaisons catégoriques	article	nom
	adjectif	nom
	déterminatif	- pronom - adjectif
	pronom personnel	- verbe - pronom + verbe
	verbe	pronom personnel
	pronom personnel postposé au	pronom
	groupes figés	
Liaisons variables	nom pluriel	- adjectif - invariable

Annexe 15 Ensemble des contextes de liaison dans lesquels seul M2 est spécifié

	M1	M2
Liaisons erratiques		<i>h aspiré</i>
		<i>oui, non</i>
		- un, huit, onze - les dérivés : <i>unième, huitième, onzième</i>
Liaisons variables		invariable
		participe passé

Annexe 16 Ensemble des contextes de liaison faisant partie de classes ouvertes

	M1	M2
Liaisons erratiques	nom singulier	
	nom propre	
	conjonction polysyllabique	
		<i>h aspiré</i>
	- adjectif singulier - pronom nasal	invariable
		- un, huit, onze - les dérivés : <i>unième, huitième, onzième</i>
	nom composé au pluriel	
	groupe figé sans liaison	
	nom pluriel	verbe
Liaisons catégoriques	article	nom
	adjectif	nom
	déterminatif	- pronom - adjectif
	pronom personnel	- verbe - pronom + verbe
	verbe	pronom personnel
	pronom personnel postposé au	pronom
	groupes figés	
Liaisons variables		invariable
	nom pluriel	- adjectif - invariable
	verbe	
		participe passé
	adverbe polysyllabique	
	adjectif qualificatif pluriel	

