

Sommaire

Sommaire	v
Avant-propos : <i>Le langage mobile du sans fil</i> , par Marc Lits	1
1. Introduction	3
1.1. De nouvelles pratiques	3
1.2. Pourquoi étudier les SMS ?	6
1.3. Une méthode scientifique	8
2. Le corpus	11
2.1. Objectif des corpus	11
2.2. Problèmes de constitution de corpus	12
2.3. Aspects juridiques	14
2.4. Récolte des données	14
2.5. Résultats en chiffres	16
3. Des données « brutes » jusqu'au corpus	19
3.1. Traitement des données brutes	19
3.1.1. Reconstitution des messages en plusieurs parties	19
3.1.2. Nettoyage du corpus	20
3.2. Anonymisation des données personnelles	21
3.3. Transcription du corpus	21
3.3.1. Pourquoi transcrire des SMS ?	22
3.3.2. Lisibilité des messages	22
3.3.3. Faciliter l'accès au contenu des SMS	23
3.3.4. Protocole de transcription	23
3.3.5. Quelques exemples de règles	24
3.4. Un logiciel de consultation	25
4. Typologie des procédés utilisés dans les SMS	31
4.1. « Phonétisation » des caractères : lettres et chiffres	31
4.1.1. Lettres	31
4.1.2. Chiffres et autres caractères	32

4.2.	Rébus	32
4.3.	Orthographe phonétique	33
4.4.	Phénomènes graphiques	37
4.4.1.	Graphies à fonction expressive	37
4.4.2.	Liaisons évoluant en agglutinations	38
4.4.3.	Graphies liées à des réminiscences	39
4.4.4.	Abréviations	39
4.5.	Icônes et symboles divers	40
4.5.1.	Smileys (émoticones)	40
4.5.2.	Symboles mathématiques et logiques	40
4.6.	Phénomènes lexicaux	41
4.6.1.	Troncations	41
4.6.2.	Sigles et acronymes	41
4.7.	Morphosyntaxe	42
4.8.	Syntaxe	43
4.9.	Discours	44
4.10.	Variété des formes	44
5.	Idées reçues	49
5.1.	Un nouveau langage ?	49
5.1.1.	Les formes lexicales	49
5.1.2.	La morphosyntaxe	53
5.1.3.	La syntaxe	53
5.2.	Le langage SMS abrège-t-il systématiquement les formes ?	55
5.3.	Existe-t-il UN langage SMS ?	57
5.3.1.	Les habitudes partagées	62
5.4.	Le langage SMS est-il un « langage oral » ?	67
6.	L'orthographe et la langue	69
6.1.	<i>Lingua non grata</i>	69
6.2.	SMS et apprentissage de l'orthographe	70
6.3.	Quand la technique s'en mêle	72
6.4.	Entrer dans les dictionnaires ?	73
7.	Quelques expériences autour des SMS	75

7.1.	15.000 milliards de phrases	75
7.2.	Un « traducteur » français > SMS	76
7.3.	Le forum	77
7.4.	Trois expériences radiophoniques	78
	Remerciements	81
	Annexe A : SMS graphiques	83
	Annexe B : Les perles	87
	Annexe C : Les mots utilisés par plus de 100 auteurs	91
	Annexe D : Préparation du corpus	95
1.	Traitement des données brutes	95
1.1.	Reconstitution des messages en plusieurs parties	95
1.2.	Élimination des doublons	96
1.3.	Élimination des parasites	96
1.4.	Élimination des messages musicaux	96
1.5.	Suppression des messages envoyés aux participants du projet	97
1.6.	Élimination du « petit poisson du bonheur »	97
1.7.	Élimination des SMS promotionnels pour des sonneries	97
1.8.	Élimination de SMS promotionnels divers	97
1.9.	Élimination de SMS adressés à l'équipe du projet	98
1.10.	Suppression des en-têtes « UCL »	98
1.11.	Recodage du symbole €	99
1.12.	Élimination des messages ïïï	99
1.13.	Rétablissement des cédilles	99
1.14.	Élimination de messages Pay&Go	99
1.15.	Anonymisation des données personnelles	99
1.16.	Mise à l'écart des SMS non français	100
1.17.	SMS graphiques	100
2.	Transcription du corpus	100
2.1.	Protocole de transcription	100
2.2.	Tags insérés dans le corpus	107
2.3.	Annotations du corpus	109

3.	Explications et mises en garde	110
3.1.	Manquement aux règles de la collecte	110
3.2.	Mots et expressions employés en français de Belgique	112
3.3.	Particularités syntaxiques rencontrées en français de Belgique	115
4.	Format final des données	116
	Bibliographie	121