

ANNEXE A

CONJONCTION	Addition	Attribution	Interrog. indirecte	Exemplification	Approximation
Le Petit Robert	?	?		0	?
Larousse	✓	✓	0	✓	✓
TLFi	?	0	0	?	0
ADVERBE	Addition	Attribution	Interrog. indirecte	Exemplification	Approximation
Le Petit Robert	?	?	✓	0	?
Larousse			0		
TLFi	?	0	0	?	0

Tableau A.1 : Emplois secondaires de *comme* mentionnés dans les dictionnaires divisés d'après le classement parmi les adverbess ou les conjonctions (✓ = mentionné/ 0 = absent/ ? = indéterminé/ gris = classé dans l'autre groupe)

CONJONCTION	Addition	Attribution	Interrog. indirecte	Exemplification	Approximation
Fr. mluvnice	0	✓	0	0	0
G. de la Phrase Fr.			0	0	0
Le bon usage	0	✓		0	?
Grammaire méth.	0	0	0	0	0
ADVERBE	Addition	Attribution	Interrog. indirecte	Exemplification	Approximation
Fr. mluvnice	0		0	0	0
G. de la Phrase Fr.	✓	✓	0	0	0
Le bon usage	0		✓	0	?
Grammaire méth.	0	0	0	0	0

Tableau A.2 : Emplois secondaires de *comme* mentionnés dans les grammaires divisés d'après le classement parmi les adverbess ou les conjonctions (✓ = mentionné/ 0 = absent/ ? = indéterminé/ gris = classé dans l'autre groupe)

ANNEXE B

sous-corpus <i>KonText</i>	N° de tokens
frWaC	1 613 814 684
L'Est républicain	146 544 420
sous-titres	32 897 232
Acquis communautaire	32 523 605
Europarl	18 968 039
core – fiction	7 194 038
PressEurop, Syndicate	6 311 555
core – nonfiction	816 023

Tableau B.1 : Taille en nombre de positions (tokens) des sous-corpus d'*InterCorp* et des corpus *L'Est républicain* et *frWaC*

sous-corpus <i>KonText</i>	f absolue	f relative	N° de types
frWaC	12 475	7,73	1740
sous-titres	735	22,34	236
L'Est républicain	549	3,75	228
core – fiction	203	28,22	120
Acquis communautaire	66	2,03	19
Europarl	28	1,48	24
PressEurop, Syndicate	20	3,17	17
core – nonfiction	1	1,23	1

Tableau B.2 : Fréquences absolues et relatives (ipm) et les nombres de types de lemmes de la construction « être + ADJ + comme » dans les sous-corpus sélectionnés de *KonText*; sous-corpus rangés dans l'ordre décroissant selon la fréquence absolue

sous-corpus <i>FRANTEXT</i>	f absolue	f relative
roman 1900 – 1997	1439	38,82
roman 1920 – 1939	451	38,71
traités, essais 1900 – 1997	287	12,37
roman 1960 – 1979	291	39,98
roman 1940 – 1959	264	37,54
roman 1980 – 1997	230	38,39
théâtre 1900 – 1997	229	66,82
roman 1900 – 1919	203	39,67
poésie 1900 – 1997	196	101,09

Tableau B.3 : Fréquences absolues et relatives de la construction « être + ADJ + comme » dans les sous-corpus sélectionnés de *FRANTEXT* rangés dans l'ordre décroissant selon la fréquence absolue

sous-corpus <i>FRANTEXT</i>	N° de textes
roman 1900 – 1919	68
roman 1920 – 1939	164
roman 1940 – 1959	82
roman 1960 – 1979	80
roman 1980 – 1997	80
poésie 1900 – 1997	111
théâtre 1900 – 1997	127
traités, essais 1900 – 1997	399

Tableau B.4 : Nombres de textes trouvés dans *FRANTEXT* pour les sous-corpus sélectionnés

ANNEXE C

rang	InterCorp logDice	f	T- score	MI	logDice	rang	InterCorp MI	f	T- score	MI	logDice
1.	roc	8	2.828	12.542	7.866	1.	barrique	4	2.000	13.750	7.007
2.	agneau	10	3.160	10.795	7.722	2.	pinson	4	2.000	13.750	7.007
3.	neige	16	3.996	9.871	7.603	3.	gardon	3	1.732	12.801	6.574
4.	linge	10	3.160	10.485	7.602	4.	berceuse	3	1.732	12.669	6.568
5.	carpe	5	2.235	11.529	7.143	5.	roc	8	2.828	12.542	7.866
6.	mule	5	2.235	10.976	7.048	6.	carpe	5	2.235	11.529	7.143
7.	barrique	4	2.000	13.750	7.007	7.	mule	5	2.235	10.976	7.048
8.	pinson	4	2.000	13.750	7.007	8.	braise	3	1.731	10.853	6.415
9.	chaton	4	1.999	10.660	6.727	9.	agneau	10	3.160	10.795	7.722
10.	gardon	3	1.732	12.801	6.574	10.	chaton	4	1.999	10.660	6.727
11.	berceuse	3	1.732	12.669	6.568	11.	linge	10	3.160	10.485	7.602
12.	plume	7	2.640	8.801	6.487	12.	pou	3	1.730	10.053	6.272
13.	pot	6	2.445	8.943	6.476	13.	neige	16	3.996	9.871	7.603
14.	clou	4	1.997	9.547	6.426	14.	chêne	3	1.730	9.867	6.229
15.	braise	3	1.731	10.853	6.415	15.	clou	4	1.997	9.547	6.426
16.	cochon	6	2.443	8.663	6.315	16.	pot	6	2.445	8.943	6.476
17.	pou	3	1.730	10.053	6.272	17.	balai	3	1.728	8.801	5.886
18.	chêne	3	1.730	9.867	6.229	18.	plume	7	2.640	8.801	6.487
19.	miel	4	1.994	8.361	5.893	19.	cochon	6	2.443	8.663	6.315
20.	balai	3	1.728	8.801	5.886	20.	miel	4	1.994	8.361	5.893
21.	glace	6	2.440	7.953	5.850	21.	planche	3	1.727	8.347	5.686
22.	enfer	6	2.439	7.865	5.787	22.	glace	6	2.440	7.953	5.850
23.	soleil	9	2.985	7.610	5.721	23.	enfer	6	2.439	7.865	5.787
24.	planche	3	1.727	8.347	5.686	24.	soleil	9	2.985	7.610	5.721
25.	chien	11	3.293	7.113	5.335	25.	chien	11	3.293	7.113	5.335
26.	bonjour	13	3.570	6.664	4.949	26.	pierre	4	1.982	6.777	4.822
27.	pierre	4	1.982	6.777	4.822	27.	bonjour	13	3.570	6.664	4.949
28.	dieu	3	1.714	6.623	4.608	28.	dieu	3	1.714	6.623	4.608
29.	roi	4	1.973	6.192	4.344	29.	ca	3	1.709	6.217	4.292
30.	ressemble r	4	1.972	6.136	4.296	30.	roi	4	1.973	6.192	4.344

31.	ca	3	1.709	6.217	4.292	31.	ressemble r	4	1.972	6.136	4.296
32.	beurre	3	1.704	5.968	4.089	32.	beurre	3	1.704	5.968	4.089
33.	air	9	2.940	5.648	3.962	33.	air	9	2.940	5.648	3.962
34.	pied	5	2.189	5.568	3.830	34.	pied	5	2.189	5.568	3.830
35.	boulot	3	1.694	5.520	3.709	35.	boulot	3	1.694	5.520	3.709
36.	bras	3	1.692	5.447	3.645	36.	bras	3	1.692	5.447	3.645
37.	mort	8	2.755	5.261	3.586	37.	client	3	1.690	5.362	3.570
38.	client	3	1.690	5.362	3.570	38.	cœur	5	2.179	5.284	3.569
39.	cœur	5	2.179	5.284	3.569	39.	mort	8	2.755	5.261	3.586

Tableau C.1 : Lemmes entrant en locutions avec *comme* et un adjectif dans la construction « être + ADJ + *comme* » dans le corpus *InterCorp* (excepté *l'Acquis communautaire*), avec les valeurs des mesures associatives *MI*, *T-score* et *logDice* affichées et les rangs ordonnés d'après *MI* ou *logDice*

rang	L'Est rép. logDice	f	T- score	MI	logDice	rang	L'Est rép. MI	f	T- score	MI	logDice
1.	gardon	9	2.999	12.277	8.157	1.	gardon	9	2.999	12.277	8.157
2.	sou	5	2.235	10.772	6.964	2.	sou	5	2.235	10.772	6.964
3.	œuf	10	3.158	9.535	6.304	3.	gosse	3	1.730	10.099	6.264
4.	gosse	3	1.730	10.099	6.264	4.	œuf	10	3.158	9.535	6.304
5.	orange	3	1.728	8.773	5.368	5.	orange	3	1.728	8.773	5.368
6.	poisson	8	2.820	8.364	5.220	6.	poisson	8	2.820	8.364	5.220
7.	pomme	4	1.990	7.623	4.458	7.	pomme	4	1.990	7.623	4.458
8.	témoigner	4	1.989	7.531	4.374	8.	Il	3	1.723	7.541	4.341
9.	Il	3	1.723	7.541	4.341	9.	témoigner	4	1.989	7.531	4.374
10.	image	8	2.805	6.937	3.866	10.	image	8	2.805	6.937	3.866
11.	pierre	3	1.716	6.728	3.600	11.	pierre	3	1.716	6.728	3.600
12.	neige	3	1.714	6.617	3.497	12.	neige	3	1.714	6.617	3.497
13.	habitude	3	1.714	6.597	3.478	13.	habitude	3	1.714	6.597	3.478
14.	Jean-Paul	3	1.713	6.531	3.416	14.	Jean-Paul	3	1.713	6.531	3.416
15.	coup	14	3.698	6.417	3.373	15.	coup	14	3.698	6.417	3.373
16.	remplacer	3	1.711	6.367	3.261	16.	remplacer	3	1.711	6.367	3.261
17.	suivre	10	3.100	5.661	2.620	17.	suivre	10	3.100	5.661	2.620
18.	gens	4	1.960	5.637	2.575	18.	gens	4	1.960	5.637	2.575
19.	cela	13	3.527	5.518	2.481	19.	cela	13	3.527	5.518	2.481
20.	monde	8	2.754	5.245	2.205	20.	monde	8	2.754	5.245	2.205

21.	la le	16	3.875	4.999	1.967	21.	la le	16	3.875	4.999	1.967
22.	celui	12	3.341	4.810	1.777	22.	celui	12	3.341	4.810	1.777
23.	chaque	6	2.358	4.749	1.709	23.	chaque	6	2.358	4.749	1.709
24.	!	6	2.290	3.938	0.904	24.	!	6	2.290	3.938	0.904
25.	dire	4	1.866	3.903	0.866	25.	dire	4	1.866	3.903	0.866
26.	un	119	10.139	3.824	0.798	26.	un	119	10.139	3.824	0.798
27.	jeu	3	1.592	3.626	0.588	27.	jeu	3	1.592	3.626	0.588
28.	trois	5	2.050	3.583	0.550	28.	trois	5	2.050	3.583	0.550
29.	on	10	2.895	3.565	0.535	29.	on	10	2.895	3.565	0.535
30.	jour	6	2.236	3.519	0.488	30.	jour	6	2.236	3.519	0.488

Tableau C.2 : Lemmes entrant en locutions avec *comme* et un adjectif dans la construction « être + ADJ + *comme* » dans le corpus *L'Est républicain*, avec les valeurs des mesures associatives *MI*, *T-score* et *logDice* affichées et les rangs ordonnés d'après *MI* ou *logDice*

rang	frWaC logDice	f	T- score	MI	logDice	rang	frWaC MI	f	T- score	MI	logDice
1.	orange	27 9	16.689	10.162	7.700	1.	bbbbrrrrrrrrrr	3	1.732	16.981	2.978
2.	agneau	34	5.818	8.797	5.639	2.	repos.de	5	2.236	16.718	3.714
3.	rétractilité	17	4.123	16.609	5.478	3.	rétractilité	17	4.123	16.609	5.478
4.	braise	20	4.468	10.139	5.472	4.	porçin	3	1.732	16.566	2.978
5.	chou	32	5.636	8.086	5.242	5.	drv	4	2.000	14.337	3.390
6.	hab	20	4.459	8.435	5.038	6.	botin	4	2.000	13.981	3.390
7.	neige	61	7.762	7.343	5.032	7.	bronchitique	3	1.732	13.866	2.975
8.	roc	15	3.867	9.447	5.010	8.	tulipa	3	1.732	13.708	2.975
9.	carpe	17	4.114	8.785	4.996	9.	Obligé	4	2.000	13.148	3.387
10.	garde-fou	13	3.603	10.265	4.944	10.	calzone	3	1.732	12.981	2.973
11.	pou	15	3.863	8.570	4.805	11.	soue	3	1.732	12.357	2.970
12.	sensation	40	6.281	7.185	4.770	12.	Beaudelaire	5	2.236	12.314	3.701
13.	habitude	60	7.677	6.809	4.589	13.	Crésus	7	2.645	12.189	4.178
14.	cochon	18	4.220	7.541	4.561	14.	coruscant	3	1.732	11.880	2.966
15.	mule	10	3.157	9.199	4.481	15.	limande	6	2.449	11.500	3.948
16.	Boîte	11	3.307	8.368	4.425	16.	chiffe	3	1.731	11.500	2.963
17.	éclair	14	3.722	7.552	4.374	17.	sylvestris	3	1.731	11.337	2.961
18.	taupe	9	2.995	9.181	4.348	18.	claustra	3	1.731	11.337	2.961
19.	bonjour	74	8.507	6.489	4.348	19.	pinson	7	2.644	10.581	4.134

20.	balai	10	3.151	8.077	4.246	20.	temperatures	3	1.731	10.538	2.948
21.	éternité	13	3.582	7.286	4.196	21.	democrates	3	1.731	10.432	2.946
22.	oeuf	44	6.556	6.424	4.194	22.	popups	3	1.731	10.328	2.944
23.	Crésus	7	2.645	12.189	4.178	23.	Ritchie	3	1.731	10.295	2.943
24.	pinson	7	2.644	10.581	4.134	24.	garde-fou	13	3.603	10.265	4.944
25.	kite	7	2.643	9.898	4.095	25.	orange	27 9	16.689	10.162	7.700
26.	témoigner	38	6.087	6.306	4.061	26.	braise	20	4.468	10.139	5.472
27.	dieu	28	5.228	6.380	4.035	27.	kite	7	2.643	9.898	4.095
28.	truc	42	6.394	6.230	4.019	28.	Crochet	4	1.998	9.731	3.324
29.	cul	12	3.437	6.977	3.987	29.	pivoine	4	1.998	9.673	3.322
30.	linge	14	3.707	6.742	3.964	30.	Sevan	3	1.730	9.618	2.922
31.	limande	6	2.449	11.500	3.948	31.	roc	15	3.867	9.447	5.010
32.	tigre	8	2.813	7.538	3.861	32.	Hérode	6	2.446	9.300	3.842
33.	Hérode	6	2.446	9.300	3.842	33.	mule	10	3.157	9.199	4.481
34.	tomate	20	4.413	6.240	3.811	34.	taupe	9	2.995	9.181	4.348
35.	planter	21	4.521	6.209	3.806	35.	Winehouse	4	1.996	9.118	3.289
36.	ornement al	6	2.443	8.508	3.751	36.	gardon	5	2.232	9.083	3.583
37.	repos.de	5	2.236	16.718	3.714	37.	pastèque	3	1.728	8.910	2.888
38.	Beaudelair e	5	2.236	12.314	3.701	38.	teigne	3	1.728	8.880	2.886
39.	papillon	10	3.131	6.655	3.694	39.	cole	3	1.728	8.833	2.883
40.	âne	9	2.973	6.811	3.690	40.	agneau	34	5.818	8.797	5.639
41.	rapace	6	2.440	7.959	3.656	41.	carpe	17	4.114	8.785	4.996
42.	amortisse ur	6	2.438	7.758	3.614	42.	ultralibéral	4	1.995	8.638	3.250
43.	camion	13	3.558	6.233	3.613	43.	pou	15	3.863	8.570	4.805
44.	mec	14	3.689	6.141	3.589	44.	fondeur	4	1.995	8.567	3.244
45.	mouchoir	6	2.437	7.628	3.584	45.	ornemental	6	2.443	8.508	3.751
46.	gardon	5	2.232	9.083	3.583	46.	rj	3	1.727	8.500	2.859
47.	singe	8	2.803	6.779	3.581	47.	hab	20	4.459	8.435	5.038
48.	ange	16	3.936	5.959	3.519	48.	linceul	3	1.727	8.410	2.852
49.	sou	8	2.798	6.538	3.473	49.	Boîte	11	3.307	8.368	4.425
50.	vôtre	8	2.797	6.513	3.461	50.	cerf-volant	4	1.994	8.266	3.211

Tableau C.3 : Lemmes entrant en locutions avec *comme* et un adjectif dans la construction « être + ADJ + *comme* » dans le corpus *frWaC*, avec les valeurs des mesures associatives *MI*, *T-score* et *logDice* affichées et les rangs ordonnés d'après *MI* ou *logDice*