

Avancement du mémoire (séance 4)

Master de Sciences du Langage

« Linguistique Théorique & Expérimentale »

Heather Burnett (LLF)

le 2 octobre, 2018

Devoir 1

Pour mardi le 9 octobre (la semaine prochaine...):

1. Choix d'un nouvel article d'où il s'agira de tirer un projet de recherche. (15 à 20 pages)
2. Résumer le travail à faire en rédigeant comme pour un(e) camarade de classe.
3. Proposer, à partir de la conclusion et des perspectives de la fin de l'article, un "projet de recherche" en 1/2 pages :
 - une question
 - une méthode
 - des résultats attendus

Pourquoi faire une étude de corpus?

Comment faire une étude de corpus?

Pourquoi faire une étude de
corpus?

Les jugements d'acceptabilité peuvent être très utiles pour investiguer des questions de grammaticalité et d'interprétation d'expressions linguistiques.

- Des fois, c'est la seule méthodologie possible...

Pourquoi faire des études de corpus?

Parce qu'on s'intéresse dans des phénomènes de **production**...

(1) Euh, Jean il n'est, euh, pas là.

Parce qu'on étudie des variétés/registres pour lesquels les locuteurs n'ont pas de jugements clairs (stigmatisation).

- (2) a. Jean n'est pas là.
- b. Jean i'est pas là.

- (3) J'omets pas mes 'ne's.

Parce que les études de corpus peuvent faire apparaître des similarités typologiques qui sont cachées par les jugements d'acceptabilité.

- (4) Toutes les personnes grammaticales en anglais peuvent être sujet/objet.
- a. I know the man.
 - b. The man knows me.

(5) Contraintes sur la voix en Lummi

- a. xč̥i-t-ŋ =sən ə cə swəy qə
know-tr-PASS =1.SING.NOM by the man
'I am known by the man.'
- b. xč̥i-t =sən cə swəy qə
know-tr =1.SING.NOM the man
'I know the man.'

Bresnan, J., Dingare, S., & Manning, C. D. (2001). Soft constraints mirror hard constraints: Voice and person in English and Lummi. In *Proceedings of the LFG01 Conference* (pp. 13-32). Stanford: CSLI Publications. E. Verhoeven.

Bresnan et al. (2001) on fait une étude de la relation entre la personne grammaticale et la voix active/passive dans le corpus *Switchboard*.

- Quand les acteurs à la première personne agissent sur des patients à la troisième personne, l'action est catégoriquement décrite avec la voix active (0/6246 occurrences).
- Quand les acteurs à la troisième personne agissent sur des patients à la première personne, l'action est décrite avec la voix passive dans 3% des cas (14/486 occurrences).

Observation

Les contraintes qui sont fortes en Lummi existent aussi en anglais, mais sont plus souples.

- Les études quantitatives peuvent révéler des similarités entre des langues typologiquement éloignées.

Comment faire une étude de
corpus?

Choisir son phénomène

Identifier des instances où deux formes expriment le même 'sens/référent (**variantes**).

- (6) a. Madame **la** ministre
b. Madame **le** ministre

- (7) a. I know the man.
b. The man is known by me.

- (8) a. Donne moi le.
b. Donne le moi.

- (9) a. J'ai vu personne.
b. J'ai **pas** vu personne.

Concordance négative

Corpus oraux

1. Anglais: Switchboard, parties du COCA...
2. Français: Corpus de français parlé parisien (CPP), Corpus ESLO

Corpus écrits

1. Anglais: BNC, parties du COCA, Google n-grams...
2. Français: Le monde, French Treebank, Frantext...

Plus c'est gros, mieux c'est (surtout pour la morphosyntaxe)...

- Les gros corpus sont souvent payants...(demandez à votre directeur.ice!)

Ressource importante pour le français

<https://www.ortolang.fr/>

Constituer son propre corpus

- Scanner des documents papier.
- Copier des pages web etc.
- Transcrire la parole spontanée (youtube etc.)

Il faut ensuite transformer les données en format xml/.txt (OCR).

- Selon ce que vous voulez faire, la constitution du corpus pourrait prendre beaucoup (ou peu) de temps.

(10) J'ai vu personne \approx J'ai pas vu personne.

Corpus oral

Le corpus *Montréal 84* de français parlé montréalais.

- 72 entretiens sociolinguistiques avec des locuteurs de genre, age et profil social variés, réalisées en 1984.
- Les entretiens sont transcrites et disponibles en format .txt.

Extraire les données

La concordance négative est possible seulement avec un sous-ensemble d'items lexicaux:

- | | | |
|------|---|-----------|
| (11) | je mangeais pas aucun legume. | (4 1526) |
| (12) | mes frères ont pas jamais été tu-vois dans le hockey | (54 149) |
| (13) | il a pas tué personne au parc Maisonneuve | (88 124) |
| (14) | Ca me tente pas de rien faire pour souper. | (118 770) |
| (15) | il-y-a pas d'anglais nulle part | (79 974) |
| (16) | j'aime pas beaucoup ni le thé ni le café. | (108 623) |

AntConc 3.4.4m (Macintosh OS X) 2014

Corpus Files
Concordance
Concordance Plot
File View
Clusters/N-Grams
Collocates
Word List
Keyword List

numero 1.txt
numero 2.txt
numero 4.txt
numero 6.txt
numero 7.txt
numero 8.txt
numero 11.txt
numero 13.txt
numero 15.txt
numero 19.txt
numero 25.txt
numero 27.txt
numero 30.txt
numero 32.txt
numero 34.txt
Numero 38.txt
numero 44.txt
Numero 46.txt
numero 49.txt
numero 50.txt
numero 51.txt
numero 52.txt
numero 54.txt
Numero 56.txt
numero 59.txt
numero 62.txt
numero 64.txt
numero 65.txt
numero 66.txt
numero 67.txt
numero 68.txt
numero 71.txt
numero 72.txt
numero 73.txt
numero 75.txt
numero 76.txt
numero 77.txt
numero 79.txt
numero 80.txt
numero 81.txt
numero 83.txt
numero 84.txt
numero 85.txt
numero 87.txt
numero 88.txt
numero 90.txt
numero 91.txt
numero 92.txt
numero 93.txt
numero 99.txt

Concordance Hits 594

Hit	KWIC
1	l\x88, <humum> puis: puis moi je suis t une
2	me d\x8Erangeait pas-mal parce- je suis une
3	" tu-sais puis: Je-veux-dire, je suis pas une
4	n anglais il parle en anglais puis je respecte la
5	cas je me dis: \x8Da peut \x90tre une:
6	\x8E des gens: veut savoir o\x90 la
7	rter de: de jugements de valeur. "Quelle sorte de
8	puis: on sait que c'est la m\x90me
9) dis \x88 elle. J: je vas (m'as) saluer la
10	' plus' de tort \x88 moi, ni \x88 toi, ni \x88
11	\x88 quelqu-un. <hum> Mais tu-sais-je-veux-dire:
12	sign\x8E. <ah oui?> () Hein? Ca c'est: t:
13	jamais pu l'avoir. Jamais. Puis il-y-a pas
14	-a bien du monde qui viennent. 2. Il-y-a pas
15	personne qui t'aide. Non. <non?> Il-y-a pas
16	e \x8Da soit' une grosse valeur pour: cette (cte)
17	uis je serai jamais parfait. Puis: je connais pas
18	vas te r\x8Epondre. <humum> Il-y-a pas
19	va m: il va: En parole il-y-a pas
20	aujourd'hui je me lave. Je suis responsable de ma
21	erreurs qui va se faire toute la journ\x8Ee,
22	8Ed\x8E, il se sentira pas au crochet de
23	: Il-y-a une chose qui ment jamais sur une
24	leur: je sais pas la personnalité\x8E d'une
25	la montagne l\x88. <oui oui oui> Je connais pas
26	8Fs trois cent soixante jours par ann\x8Ee, puis
27	nulle part. Tu-sais: Le bouffon il-y-a jamais
28	Puis c'est des effusions qui tu entends pas \x88

Search Term Words Case Regex

Search Window Size

personne | Advanced | 50

Start Stop Sort

Kwic Sort

Level 1 1R Level 2 2R Level 3 3R

Total No. 83

Files Processed

- Exporter les données dans un tableur.
- Éliminer les exemples qui ne sont pas pertinents.

- (17) a. la/une personne
b. Personne est venu \nrightarrow *Personne est pas venu.

Codification

avm_example.csv

Liberation Sans 10

C7 f_x Σ = pas l'ouvrage que je fais. Je reparerais aucun équipement électronique tandis-que

	A	B	C	E	N	O	Q	S	U
1	Locuteur	ligne	exemple	concordance	genre	age	scolarite	prof.84	indefini
2		1868	runer tout'. Il veut: il veut: il accepte aucun: compromis. Il accepte aucune	b	M	38	1		6 aucun
3		2026	Michelle coeur de punk (rire) 2. Aucun: CFG2. 2. Je suis pas capable. 2.	b	M	38	1		6 aucun
4		1290	Franchement il-y-a pas: il-y-a pas aucun: l'écoute pas un programme je pense avec pas	b	F	31	1		4 aucun
5		1398	pas là. En tout cas on fait partie d' aucun de t ca. Mon marj aime pas les	b	F	31	1		4 aucun
6		1526	en: je mange pas: je mangeais pas aucun legume. Fait-que là les	b	F	31	1		4 aucun
7		160	pas l'ouvrage que je fais. Je reparerais aucun équipement électronique tandis-que	b	M	38	2		3 aucun
8		25	jours possibilité parce-que il-y-a aucun empechement mais disons que s: il	b	F	37	2		4 aucun
9		381	oui. D-abord on n'a pas de meeting, aucun meeting, parce-que ça se fait à	b	F	37	2		4 aucun
10		1084	2. Pas du tout. 2. J'en achete aucun. On a les journaux locaux ici	b	F	37	2		4 aucun
11		1162	à la maison fait-que: j'avais pas aucun temps pour moi. (#G7) <2. Non><2. Non>	b	F	37	2		4 aucun
12		172	Pas de: J'ai au: j'ai aucun métier. Je suis chauffeur de	b	M	41	1		5 aucun
13		367	2. Ah bien oui. Moi: Je suis pas aucun programme anyway: Les programmes qu'ils	b	M	41	1		5 aucun
14		108	ça la. Tu-sais: Parce-qu'il j'avais pas aucun truck à part de t ca qui pouvait	b	M	59	2		6 aucun
15		898	part ça a pas l'air apparent. Ca n'a aucun lien mais: () en tout cas il-y-a une	b	F	37	3		4 aucun
16		925	mais moi les voisins je frequente pas aucun voisin. "Bonjour bonjour." Même	b	F	72	1		4 aucun
17		1669	aussi beau. Parce-qu'elle faisait aucun geste. La: la deuxième fois qu'on a	b	F	72	1		4 aucun
18		1727	c'est bon. Mais moi j'aime aucun vin. Non. (#G11) 2. Quelque chose	b	F	72	1		4 aucun
19		370	système an: francophone ils apprenaient aucun anglais avant la quatrième année.	b	F	29	3		2 aucun
20		1672	2. Non. 2. Non j'ai pas ca, j'ai aucun de ca. (rire) 2. Lequel je	b	F	39	1		4 aucun
21		72	du temps. (#S6) 2. Non. 2. Aucun cours. 2. Non. 2. Saturé. Fini.	b	M	36	2		3 aucun
22		76	natation mais à part de t ca là, aucun: Oh non. Elle en a assez de l'ecole ma	b	M	36	2		3 aucun
23		311	Non ça dépend, ça dépend. il-y-a aucun poste que je peux dire que j'aime	b	M	36	2		3 aucun
24		335	(#G2) 2. Moi, non. Du tout. 2. Non. AUCUn (#G3) 2. Tous les jours. 2.	b	M	36	2		3 aucun
25		62	179 qu'il: Parce-que là, disons c'est aucun: J'ai pas pense bien bien à rien	b	M	30	1		3 aucun
26		243	4 pied tranquillement puis: il-y-a aucun problème. M'me Si tu as pas de laveuse	b	M	40	3		3 aucun
27		432	l'interieur comme de l'exterieur il-y-a aucun problème là-dessus. Ça aurait été	b	M	40	3		3 aucun
28		494	un blocage peut-etre: Quand j'ai: aucun système de chauffage sauf 'système au	b	M	40	3		3 aucun
29		365	de la place pour un bon homme ca c'est: aucun doute là-dessus. 2. (grande	b	M	64	3		1 aucun
30		386	brillants, ils sont intelligents. Il-y-a aucun doute là-dessus. Le seul point qui	b	M	64	3		1 aucun
31		394	qualités, connaissances ca: aucun doute là-dessus. 2. Mais ça: ça	b	M	64	3		1 aucun
32		655	ça ça s'exprime dans le vêtement, aucun doute. Ca s'exprime: Deux gars ont une	b	M	64	3		1 aucun
33		786	danseuses de ballet etcetera. Ca: j'ai aucun objection contre ça. Mais de dire	b	M	64	3		1 aucun
34		24	problème pour moi. 2. Ah oui () aucun problème. (#R4) 2. Oh: c'est le	b	M	70	3		1 aucun
35		415	pas besoin du: du titre oui () aucun problème. J'aurais suivi ça en simple	b	M	70	3		1 aucun
36		933	aucun que je suis régulièrement, aucun Non. 2. Bien certains programmes	b	M	70	3		1 aucun
37		1028	la musique, c'est très beau, il-y-a aucun doute, mais moi je connais rien a la	b	M	70	3		1 aucun
38		48	étudié alors: 2. Jusqu'à date aucun problème. (#O2) 2. Pas pour le moment	b	M	30	3		2 aucun
39		244	pas à ca: du tout. Pour moi ca a aucun rapport () peut-etre parce: je suis	b	M	30	3		2 aucun

Résultats: Indéfini

Indéfini	Sans pas	Avec pas	Total	% pas
Nulle part	3	15	18	83%
Personne	49	71	120	59%
Rien	800	137	937	15%
Aucun	134	16	150	11%
Jamais	1134	11	1145	1%
Total	2120	250	2370	10.5%

Résultats: Sclarité

Niveau d'éducation	Sans pas	Avec pas	Total	% pas
Bas	643	114	757	15%
Moyen	634	80	714	11%
En cours	139	12	151	8%
Élevée	683	44	727	6%

Pour évaluer la signifiante de ces pourcentages, il faut faire des analyses statistiques...