**Categorical Statistics for Linguistics**

*Méthodes pour les débutsants / Introductory Textbooks*

To date, there are not any text books for statistics for linguists in French, here are a selection of the most popular international textbooks

**Baayen (2008)**

Harald Baayen is a Dutch grammarian and he arguably the most respected linguist statistician today. Put simply, everyone else learnt their basics from him and he incredibly respected by all concerned. This is a veritable bible, but quite difficult to follow. This book will be important for the M2 group, but students in M1 are strongly recommended to consult it as well.

Chapters – 1, 2, 3, 4, 5, 6, 7 (all the chapters)


**Dalgaard (2008)**

Peter Dalgaard is a Danish statistician. His perspective is different from ours, but his statistical knowledge is very up to date and trustworthy. An excellent book with clear explanations, I strongly recommend that everyone consults it.

Chapters – 1, 2, 4, 6, 8, 13


**Johnson (2008)**

Keith Johnson is an American phonetician, so his take on the subject is quite different from ours. A book that is structured oddly, but is written clearly. Several of the chapters are very useful.

Chapters – 5, 6, 7


**Glynn & Robinson (2014)**

Justyna Robinson is a Polish historical linguist, her and I put this book together to bridge the gap between the snobbery of the quantitative linguistic community and the rest of the linguistic community. It is a book we will use very closely, both at M1 level and M2 level, but esp. M1 level. It has a chapters on several statistical techniques written for beginners as well as a selection of articles which give examples of their use.

Chapters – Section 1


**Gries (2009)**

Stefen Gries is a German corpus linguist and a machine. People say he writes quicker than most people can read. He is extremely prolific and extremely demanding of his students. Many people see him as the next Harald Baayen. This book is very popular but not everyone likes it. There is a new version (2013) which is greatly expanded and arguably much better.

Chapters – 2, 4, 5 (Gries 2013)

Oakes (1998)
In my opinion, this is a disappointing book. Mainly because I feel it focuses on things that a debutant does not need to know. But, it certainly has some excellent explanations of “basic” concepts


Woods et al. (1986)
I only mention this book because it is so common and is in most libraries. It was an excellent book in its day but modern computers have rendered most of its contents outdated. However, many of the basic concepts are very well explained here.


Gómez (2013)
Pascual Cantos Gómez is a Spanish computational linguist. This book is quite new. Most of it designed especially for corpus linguists, but chapter 4 is directly relevant to us.


Lowie & Seton (2012)
Wander Lowie and Bregtje Seton two Dutch applied linguists, I found this book disappointing and certainly relevant to our course.


Larson-Hall (2009)
Jenifer Larson-Hall is an American applied linguist. This book is not very helpful in this course.