Cómo realizar un análisis estadístico multivariado en Rbrul con datos numéricos/continuos

Introducción

Este manual sirve como guía para estudiantes que quiere aprender a realizar análisis estadísticos en Rbrul. Trabajaremos específicamente con datos numéricos o continuos¹ ya que existen varios manuales que se enfoca en datos categóricos (véase <u>http://www.danielezrajohnson.com/rbrul.html</u>). Rbrul tiene la capacidad de realizar análisis de regresión, el cual es un proceso estadístico que evalúa las relaciones entre variables. El análisis de regresión permite observar cómo el valor de una variable dependiente cambia cuando cualquiera de las variables independientes es variada, mientras que se mantienen fijas las otras variables independientes. Hay regresiones logísticas, las cuales se realizan con datos binarios, y por lo tanto, los datos se expresan en proporciones o porcentajes. También hay regresiones lineales, las cuales se realizan con datos numéricos, y por lo tanto, los datos se expresan como promedios. Nuestro enfoque será la última.

Terminología

Antes de continuar, es necesario aclarar algunos términos. Una **variable** es una propiedad que tiene valores distintos. Es decir que varía. Por ejemplo, la /s/ española en posición de coda es una propiedad lingüística que puede ser pronunciada de formas diferentes, una sibilante, una aspirada o una forma elidida. Otra propiedad variable que se encuentra en cualquier lengua es acentuación silábica. Generalmente hablando, hay sílabas átonas y sílabas tónicas.

En las estadísticas, existen **variables dependientes** y **variables independientes**. Una variable dependiente es aquella que responde a los valores de otra variable. Es decir, los valores de la variable dependiente *dependen* de los valores de otra variable. También es conocida como *variable respuesta*. Los investigadores de la sociolingüística cuantitativa, por lo general, hablan de una *variable lingüística*, la cual viene siendo una variable dependiente; por ejemplo, la /s/ española en posición de coda. Recuerde que es la variable bajo estudio. Los valores de la variable dependiente (o respuesta o lingüística) son conocidos como **variantes**.

Las variables independientes son aquellas cuyos valores influyen en la variable dependiente y, por lo general, no dependen de los valores de otras variables. También se conocen como *variables explicativas* porque *explican* la variación de la variable dependiente. Por

¹ Los datos utilizados en este manual viene de un corpus de español panameño que contiene hablantes monolingües y bilingües. Se trata de un análisis de la producción de la /s/ en posición de coda en términos de una propiedad acústica conocida como centro de gravedad (o *center of gravity* en inglés). El centro de gravedad (CDG) es un rasgo acústico de sonidos fricativos. Mide la energía acumulada en la fricación. Mientras mayor el centro de gravedad, más energía concentrada tiene el sonido. En el caso de la /s/, se usa para detectar la reducción o debilitamiento de dicho sonido. Los sonidos fricativos linguales que tienen un punto de articulación anterior, como por ejemplo la sibilante, tienden a tener mucha energía en la fricación, y por lo tanto, un centro de gravedad alto. Los sonidos fricativos linguales que tienen un punto de articulación posterior, como la aspirada glotal (una forma reducida de la /s/), tienden a tener poca energía de la fricación, y por lo tanto, un centro de gravedad alto, un centro de gravedad bajo (véase File-Muriel & Brown, 2011 para más información sobre esta propiedad y otras relacionadas a la /s/).

ejemplo, una sílaba átona tiende a favorecer el uso de una [s] aspirada mientras que una sílaba tónica tiende a favorecer el uso de una sibilante. La acentuación silábica es una variable independiente que explica la variabilidad de la propiedad lingüística /s/. En la tradición de la sociolingüística cuantitativa, las variables independientes son conocidos como *grupos factoriales* porque cada uno de sus valores es un factor, o sea que es un grupo de factores. Veremos que en Rbrul se usa el término *predictor* para referirse a factor ya que el predictor *predice* el comportamiento de la variable dependiente. En lo que sigue, se utilizarán indistintamente los términos variables independientes y grupos factoriales.

1. Abra R

2. INSTALAR PAQUETES

Por lo general, este paso sólo es necesario la primera vez: En el menú principal, seleccione "Package & Data" y dele clic a "Package Installer". Se abrirá una nueva ventana. Esta ventana permite hacer una búsqueda de todos los paquetes disponibles en el mundo o ambiente de R.

🗯 R File	Edit Format Workspace F	Packages & Data Mis	sc Window Help			Q & \$ 🕫 🚯 🕴	奈 🜒) 🔲 100% 🖾	Mon 6:56 PM	২ ≔
000		_		1	R Console				
🥶 🦃 🏦	1 🛍 🚞 🌖 🖄 📃								<u>_</u> B
[~		T MEENING.					Q Help Search		
Natural lan		R Package Installer							
R is a collab	Packages Repository								
Type 'contrib 'citation()'	CRAN (binaries)	0							
Type 'demo()' 'help.start() Type 'q()' to	Get List 🗸 Binar	y Format Packages	Q-Package Search	8					
[R.app GUI 1.	Package	Installed Version	Repository Ve	rsion					
[Workspace re [History rest									
> source("htt									
> rbrul()									
Welcome to R Your current									
In case of t Please don't									
Windows user or when upda									
If you get a are in your									
Rbrul is no converge pro									
Successfully Content type					'MuMIn_1.14.0.tgz'				
downloaded 22	Install Location								
uomittoudeu 33	• At System Level (in R framework)		Install Selec	ted					
The downloade	At User Level		Install Depender	cies					
Successfully	 In Other Location (Will Be Asked U As defined by .libPaths() 	Ipon Installation)	Update A						
No data loadea									
MAIN MENU									
1-load/save date 9-reset 0-exit	1								
1:									

En el primer menú de esa ventana, debe estar seleccionado "CRAN (binaries)". Luego, dele clic a "Get List". Se llenará la lista de paquetes.

•						
Packages Repository						
CRAN (binaries)	2					
Get List 🗸 B	Q-Package Search					
Package	Installed Version	Repository Version				
A3		0.9.2				
abc		2.1				
abc.data		1.0				
ABCanalysis		1.0.2				
abcdeFBA		0.4				
ABCoptim		0.13.11				
ABCp2		1.1				
abctools		1.0.2				
abd	0.2-7 0.7-1 1.4-3					
abf2						
abind						
abn	0.85					
abundant		1.0				
acc		1.0.2				
accelerometry		2.2.5				
AcceptanceSampling		1.0-3				
ACCLMA		1.0				
accrual		1.0				
Install Location						
• At System Level (in R framewo	Install Selected					
At User Level		Install Dependencies				
 In Other Location (Will Be Ask 	ed Upon Installation)					
 As defined by .libPaths() 		Update All				

A la derecha de "Get List", verá una barra de búsqueda donde escribirá **Ime4**. Este paquete permite hacer análisis estadísticos multivariados (y otras cositas divertidas) en Rbrul. Al escribir el nombre del paquete, verá que se reducirá la lista y aparecerá el paquete necesario. Si ya está instalado, la segunda columna indicará la última versión instalada. La tercera columna indicará la versión actual disponible. Si no está instalado (o si tiene una versión antigua), instálelo seleccionando el paquete en la lista y dándole clic a "Install Selected", el cual se encuentra por debajo de la lista de paquetes. RECOMENDACIÓN: ponga un gancho ✓ en el cuadrito de "Install Dependencies". Esta opción permite instalar paquetes adicionales de los cuales depende el paquete Ime4. Algunos son absolutamente necesarios y otros no. Si no le pone gancho, tendrá que instalar estos paquetes manualmente. Después de la instalación, puede cerrar la ventana.

	R Package Installer	
Packages Repository		
CRAN (binaries)	٥	
Get List	Binary Format Packages	Q Ime4 🚳
Package	Installed Versio	n Repository Version
lme4	1.1-7	1.1-7
Install Location		
At System Level (in R	framework)	Install Selected
At User Level In Other Location (Wi	Be Asked Upon Installation)	Install Dependencies
	-0	Lindate All

3. CARGAR RBRUL

El programa Rbrul es un archivo de texto con el nombre "Rbrul.R". Para usarlo, tiene que cargarlo en R. Hay dos maneras de hacerlo:

PRODECIMIENTO A

Puede usar el siguiente comando:

source("http://www.danielezrajohnson.com/Rbrul.R"). Cópielo y péguelo en la ventana principal de R. Después de un momento, aparecerá una flecha ">" debajo del comando. **OJO**: Este procedimiento requiere una conexión al internet ya que estará accediendo al programa desde la página web de Dan Johnson.



PROCEDIMIENTO B

Para acceder al programa sin internet, es necesario guardar el programa en su computadora. Diríjase a este enlace:

<u>http://www.danielezrajohnson.com/rbrul.html</u>. Allí se encuentra una lista de enlaces. El primero es el enlace a la última versión del programa Rbrul.

RBRUL by Daniel Ezra Johnson

"A tool such as Rbrul offers a compromise of the old and new that I believe will be widely used in the near future." - R. Harald Baayen

"Using mixed models and adding individual speaker as a random effect results in interesting, logical results for my data. The results are conservative, but I like that. If I don't use speaker as random, I get loads of extra factors as significant, but lots of these make no sense and simply can't be explained. This again gives me confidence in my conservative approach." – a satisfied customer

"I've been using it a lot and finding it so much easier than trying to do the same in R." - another satisfied customer

"Thanks Dan. You cast light upon my path into statistical wisdom." - another real testimonial

"I really like how you designed Rbrul. It is so user-friendly! And fast, too! Perhaps the best thing about using Rbrul is Dan's speedy responses to questions I pose to him. Thank you, Dan!" - more not-made-up feedback



Haga clic derecho en ese enlace y seleccione "Save Link As…" para guardar el programa como archivo de texto en su computadora. El nombre del archivo debe ser "Rbrul.R". **OJO**: Dan Johnson actualiza regularmente el programa, por lo que tendrá que volver a guardar el archivo periódicamente para asegurarse de que tenga la versión más actualizada. Se recomienda editar el nombre del archivo según el día en que fue guardado en su computadora, por ejemplo, "Rbrul_06222015.R". Esto le recuerda de cuándo fue la última vez que guardó el archivo. También, si por alguna razón no funciona la versión actualizada en su computadora, puede volver a usar la versión anterior.

Para cargar el programa en R, en el menú de "File", dele clic a "Source file", busque el archivo de texto "Rbrul.R" (o cómo lo haya denominado) y selecciónelo.

			R Console				
2	🕯 🗅 🔚 🜖 🖺	8					
1	000		R File to Source		<u></u>		
on 3.1.3			Searching "This Mac"	\$	Q, rbrul.r		٥
Platform: x86_€	Favorites	Search: This Mac "New Analysis_05272015"					Save +
ee softw	Recents	Documents	^	Date Modified	Date Created	Size	Kind
icense()	Crophox	RandomForestBinary.txt		Oct 29, 2014, 8:07 PM	Oct 29, 2014, 8:07 PM	368 KB	Plain Text
al langı	All My Files	PDF Documents					
-ollohar	All my riles	drummond 2012.pdf		Feb 2, 2014, 12:11 PM	Feb 2, 2014, 12:11 PM	278 KB	Adobeume
ontribut	Cloud Drive	Getting_started_with_Rbrul.pdf		Jan 22, 2014, 2:51 PM	Jan 22, 2014, 2:51 PM	2.2 MB	PDF Docum
on()' or	Applications	Michnowicz & Carpenter 2013.pdf		Oct 15, 2014, 2:18 PM	Oct 15, 2014, 2:18 PM	1.7 MB	Adobeume
eno()' f	Desktop	1 rbrul handout_daleszynska.pdf		Mar 29, 2012, 8:24 AM	Mar 29, 2012, 8:24 AM	624 KB	PDF Docume
tart()'	Documents	Developer					
	O Deverte ede	Rbrul_219.R		Sep 10, 2014, 12:05 AM	Sep 10, 2014, 12:05 AM	102 KB	R Source File
GUI 1.65	Downloads	Rbrul.R		Aug 12, 2014, 5:03 AM	Aug 12, 2014, 5:03 AM	106 KB	R Source Fill
ace rest	Google Drive	Rbrul3302015.R		Mar 30, 2015, 11:35 PM	Mar 30, 2015, 11:35 PM	108 KB	R Source File
/ restor	Devices	Rbrul05062015.R		May 8, 2015, 9:46 PM	May 8, 2015, 9:46 PM	109 KB	R Source File
	Remote Disc						
	MICROSOFTO ≜						
	Prest64 5409						
	□ Prost22 5261 ▲						
	Shared						
	📕 user-pc						
	Madia	delanosyd ► Documents ► Rbrul05082015.R					
						Cancel	Open
						Ganger	open

4. RBRUL

Ahora puede correr Rbrul escribiendo el comando **rbrul**() en R y dándole a "Enter". Después de un momento, aparecerá una descripción del programa. También se cargarán los paquetes necesarios para realizar los análisis estadísticos multivariados.



El más importante es el paquete **lme4**. Si los paquetes no están actualizados, se actualizarán automáticamente y luego se cargarán. Por último, aparecerá el menú principal de Rbrul.

Console	K File Edit Format Workspace Packages & Data Misc Window Help	🗛 🕸 🕙 🗍 🤝 🖜) 📼 100% 📾 Mon 6:50 P	M Q ≔
 We way and the second se	Console		
	😑 🕀 🛍 🏠 🚍 🔾 🖺 🗋 🔒		[8]
<pre>ktural language support but running in an English locale B is a collaborative project with many contributors. Pye is contributors() for nose integers in positivations. Pye is contributors() for nose it is a for English locale integers in positivations. Pye is contributors() for nose it is a for English locale integers in positivations. Pye is contributors() for nose it is a for English locale integers integers in positivations. Pye is contributors() for nose it is a for English locale Pye is contributors() for nose it is a for English locale Pye is contributors() for nose it is a for English locale integers Pye is contributors() for nose it is a for English locale integers Pye is contributors() Pye is contributo</pre>	[=	Q Help Search	
B is a collaborative project with many contributors. Dye identification? Dy	Natural language support but running in an English locale		
<pre>pype ieenO; for some denos, 'NelPO', for on-line help, or 'ype igO' to guit 8. (Ropp GUI 1.65 (GB1) x66.ch-ople-domini1.4.0] (Norkspace restored from /Users/delamosyd/.Ropt.iterO]) issuer ("http://www.dmileizerojonson.comRbml.8") > source("http://www.dmileizerojonson.comRbml.8") > source("http://www.dmileizerojonson.comRbml.8") > source("http://www.dmileizerojonson.comRbml.8") > source("http://www.dmileizerojonson.comRbml.8") > source("http://www.dmileizerojonson.comRbml.8") > source("http://www.dmileizerojonson.comRbml.8") > source("http://www.dmileizerojonson.comRbml.8") Nelows sure how sure your 6 version is updated, and try loading Bbml again. Pleese don't housing source is to updated, and try loading Bbml again. Pleese don't housing source is updated, and try loading Bbml again. Pleese don't housing source is to updated, and try loading Bbml again. If you get an error soying 'invalid multibyte string', it mens foreign characters are in your durities to domsels, or even worse is et vid of them! Successfully loaded lear package.trying UL 'http://stit.eths.ck/CMAVbin/noccess/movericks/contrib/3.1/AddIn_1.14.0.tgr' commlanded binary packages are in</pre>	R is a collaborative project with many contributors. Type 'contributors()' for more information and 'citation()' on how to cite R or R packages in publications.		
<pre>[R.app GUI 1.65 (6913) x86.64-apple-damin13.4.9] [Dirkspor restored fram /Users/delmosyd/.Botto] [Dirkspor restored fram</pre>	Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help. Type 'd()' to quit R.		
<pre>[privacyses restored from /users/delmosyd/.Botel] [history restored from /users/delmosyd/.Botel] [histor /users/delmosyd/.Botel] [</pre>	[R.app GUI 1.65 (6913) x86_64-apple-darwin13.4.0]		
<pre>> source("http://www.domlelezrgjohnson.com/Rbrul.R") > for com/Rbrul.R") Retcome to Drul version 2.28 - High Beams Your current R version 15 3.1.3 - Smooth Sidewalk In case of frouble, make surve your R version is udotted, and try loading Rbrul again. Please don't hesistate to email donielerrgjohnson@mail.com for help. (This works.) Windows users may need to run R as administrator the first time they use Rbrul, or men updates cource. Right-clifts of administrator the first time they use Rbrul, or men updates cource. Right-clifts, or even workel Ext right them! Rbrul is no longer suppressing (red) convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't converge roperly may be too convergence wornings. Models that didn't convergence wornings. Models that your am risk. Successfully loaded ling roperly may be too convergence wornings. Wind KMU MeMU MeMU -load/awe data -lo</pre>	[Workspace restored from /Users/delanosyd/.RData] [History restored from /Users/delanosyd/.Rapp.history]		
Perul() Welcome to Khrul version 12.32 - High Beams Your current R version 13.3.1 - S south Sidewalk In case of trouble, mole sure your R version is updated, and try loading Rbrul agpin. Please don't heistate to email dividerary ofhoms/molic. On for help. (This works.) Nindows users may meed to run R as administrator the first time they use Rbrul, or when updates occur. Right-Click to do this. You may need on admin password. If you get an error soying 'involid multibyte string', it means foreign chonceters are in your data: accents, diacritics, diarchies, or even worsel Get rid of them! Rould so and the data concert solution works and the theorem works is defined to the source solution (red) complex for the data. Trut them at your ann risk. Successfully loaded heat package. commonded J37 K8 The downloaded binary packages are in /var/folkers/nk/sjdg22corijs/my_Livet.min/sites/downloaded_packages Successfully loaded Multin package. Successfully loaded Multin package. Successfully loaded Multin package. Nin Meku -load/ave data -load/ave data	<pre>> source("http://www.danielezrajohnson.com/Rbrul.R")</pre>		
Relcome to Bbrul version 3.28 - High Bems Your current R version 13.1.3 - Smooth Sidemalk In case of truble, make sure your R version is updated, and try loading Rbrul appin. Please don't hesitate to mail danielezrajohnson@publication for help. (This morks.) Windows users soy need to run R as administrator the first time they use Rbrul, or men or soying 'involtantitotic durity best string', it means foreign characters are in your data: accents, diacritics, diagraphs, or even mores! Get rid of them! Rbul is no longer suppressing (rev) convergence wornings. Model: that dian't converge property may be too complex for the data. Truit them at your own risk. Successfully loade lamed package.trying IRL 'http://statiethc.d/UGM/bin/macoss/movericks/contrib/3.1/MMIn_1.14.0.tgz' comend URL' dominaded binary packages are in /var/folders/ms/Bgdd2crijsdew_1///RtmpnfMWKp/domloaded_packages Successfully loaded binary packages are in /var/folders/ms/Bgdd2crijsdew_1///RtmpnfMKp/dowloaded_packages Successfully lo	> rbrul()		
In case of trouble, mole sure your R version is updated, and try loading Bkrul again. Please don't hesistate to email donielerzio/showedmomphil.com for heb. (This works.) Windows users may med to run R as administrator the first time they use Bkrul, or men updates occur. Right-click to do this. You may need an admin passmord. If you get an error soying 'invold multiply esting,'it means foreign characters are in your data: accents, diacritics, diarophs, or even morsel Get rid of them! Rowl is no longer suppressing (red) convergence marrings. Nodels that dian't converge properly may be too complex for the data. Trust them at your am risk. Successfully loaded heat package. Comtonit type 'emplicition/x-tar' length 346922 bytes (337 KB) The dowlloaded binary packages are in /var/folkers/nk/sjdd2chrijsdewggn/T//RinpnfMWkp/downloaded_packages Successfully loaded Multi package. Noded Multi package. No data loaded. Min MeMU -load/awe data	Welcome to Rbrul version 2.28 - High Beams Your current R version is 3.1.3 - Smooth Sidewalk		
Windows users way need to run 8 as administrator the first time they use Borul, or metru pddstock.rkliket.rkliket to borul in your own admin passmod. If you get an error saying 'involid multibyte string', it means foreign characters are in your data: accents, diacritics, digraphs, or even moreal feat rid of them! Borul is no longer suppressing (red) convergence warnings. Models that didn't converge properly may be too complex for the data. Trust them at your own risk. Successfully loaded hard package trying UM: "http://tstr.etw.c.w/CMAVbin/maccoss/movericks/contrib/3.1/AAMIn_1.14.0.tgr' complex for the data. content type 'opplication/x-tar' length 346022 bytes (337 KB) ownHoudded binary packages are in // Arr/folers/n/63/680080pn/T//RtmpnMMXp/downloaded_packages Successfully loaded Muth package. Num MNU MNU -load/axee data	In case of trouble, make sure your R version is updated, and try loading Rbrul again. Please don't hesitate to email danielezrajohnson@gmail.com for help. (This works.)		
If you get an error soging 'involid "ultibyte string', it means foreign characters are in your data: accests, discritics, digraphs, or even worsel Get rid of then! Rbrul is no longer suppressing (red) convergence warnings. Models that didn't converge properly may be too complex for the data. Trust them at your and risk. Successfully loaded heet packages trying UK: "http://stat.eth.c.//CAN/bin/macoss/movericks/contrib/3.1/MAMIn_1.14.0.tgz' content type 'application/x-tar' length 346822 bytes (337 KB) converloaded binary packages are in /war/folkers/mc/Sgd2Ecr5jgmy_Liwfanjs8000gn/T//RtmpnMMKp/downloaded_packages Successfully loaded MAIn package. Successfully loaded MAIn package. Successfully loaded MAIn package.	Windows users may need to run R as administrator the first time they use Rbrul, or when updates occur. Right-click to do this. You may need an admin password.		
Rbnl is no longer suppressing (red) convergence wornings. Models that didn't converge property may be too complex for the data. Trust them at your am risk. Successfully loaded late package. trying UL "http://isti.eth.eth.ch/CAAVbin/mccoss/movericks/contrib/3.1/AAMIn_1.14.0.tgt' content type "upplication/x-tar" length 346822 bytes (337 KB) content type "upplication/x-tar" the downloaded binary packages are in //war/folder.n/M5/gdd2cdrigdety_Liveqnjb8000gn/T//RtmprfMKG/downloaded_packages Successfully loaded MMEIn package. Successfully loaded MMEIn package. No dato Loaded. MAIN MBNU -Inadd/sowe data	If you get an error saying 'invalid multibyte string', it means foreign characters are in your data: accents, diacritics, digraphs, or even worse! Get rid of them!		
Successfully loaded lame package.trying UBL "http://iot.etxc.dvCRAW/bin/mocoss/movericks/contrib/3.l/AAMIn_1.14.0.tg" content type 'application/x-tar' length 346622 byte (Internet Total Content Type 'application's-tar' length 346622 byte (Internet Total Content Type 'application's application's application'	Rbrul is no longer suppressing (red) convergence warnings. Models that didn't converge properly may be too complex for the data. Trust them at your own risk.		
downloaded 337 K8 The downloaded binary packages are in /var/folkers/nk/sjdgd2crijjeky_lwe_in/sp8000gn/T//RtmpnfMWXp/downloaded_packages Successfully loaded MAIT package. No data loaded. MIN MENU 1-load/ave data 3-rest d ==st	Successfully loaded lme4 package.trying URL 'http://stat.ethz.ch/CRAW/bin/macosx/mavericks/contrib/3.1/MuMIn_1.14.0.tgz' Content type 'aplication/x-tar' length 346022 bytes (337 KB) opened URL		
The downloaded binary packages are in //ar/folders/n/8/5gd2c6n5jdetw_t_lwenjp8000gn/T//RtmpnfMWXp/downloaded_packages Successfully loaded MAIT package. No data loaded. MAIT MENU 1-load/arwe data 3-raget 0-exit	downloaded 337 KB		
Mine Doministration United publications of the fit in the Constant of the Cons	The downloaded himes, performe and in		
Successfully loaded Mattice package. Successfully loaded lattice package. Main MENU A regist B -regist B -regist B -regist B -regist B - regist	/var/folders/n8/5g6d26cn5jg4myt_lw4qnjp80000gn/T//RtmpnfMWXp/downloaded_packages		
No data loaded. MLTN HENN 9-reset 0-exit 1	Successfully loaded MuMIn package. Successfully loaded lattice package.		
MATN HENN 1-load/sowe data 9-reset 0-exit	No data loaded.		
	MIN KENI 1-load/swe data 9-reset 8-exit 1:		

5. CARGAR LOS DATOS

rbrui(packages = T)

Escriba "1" para cargar o leer los datos que se analizarán y dele a "Enter". El archivo debe estar en formato .csv o "comma-separated values".

No data loaded.

MAIN MENU 1-load/save data 9-reset 0-exit 1: 1

Luego, escriba "c" y dele a "Enter".

MAIN MENU 1-load/save data 9-reset 0-exit 1: 1

No data loaded.

What separates the columns in the data file to open? (c-commas s-semicolons t-tabs tf-token file) Press Enter to exit, keeping current data file, if any. 1: c

Se abrirá una ventana para seleccionar el archivo de datos.

If Console Choose File It does to console It does to console If will 1.65 (Gill) If will 1.65 (Gill)<
Image: Section 1. Im
Image: Superior Control Choose File Q Help Sueuch View (a) View out it &. Choose File Q Help Sueuch View (a) View out it &. Choose File Q Search View (a) View out it &. Choose File Q Search View (a) View out it &. Choose File Q Search Periopsoce restored from /Users/delamosyd/.Sapp.history Sum 8, 2016, 8,21 FM Aun 8, 2016, 8,21 FM Aun 9, 2016, 8,01 FM > source("http://www.dnail@craigheson.com/Binul.F") A IN My Files Popbox Autor Autor Choose File Dooments Autor Autor Choose File Autor Autor Choose File * Choose File Counters StantactVariability/2015 Aun 9, 2016, 8,01 FM Aun 9, 2016, 8,01 FM Aun 9, 2016, 8,01 FM * StantactVariability/Standows Aun 9, 2016, 8,01 FM Aun 9, 2016, 8,01 FM Aun 9, 2016, 8,01 FM Autor 2, 2016, 8,27 AM * on the updates occurs, Night-Click to do this. You may if the stantact it and updates occurs, Night-Click to do this. You may if the outperiod choose file for the stantact it and updates occurs, Night-Click to do this. You may if the outperiod choose file for the stantact. Oberlaw, Augustatt Auror 2, 2016, 8,27 M Auro, 2016, 1102 M Auro, 2016, 1028 M Au
Image: State of the and ML provide interface to help: Up: A part of the and ML provide interface to help: Up: A part of the and ML provide interface to help: Image: State of the and ML provide interface to help: Image: State of the and ML provide interface to help: Image: State of the and ML provide interface to help: Image: State of the and ML provide interface to help: Image: State of the and ML provide interface to help: Image: State of the and ML provide interface to help: Image: State of the and ML provide interface to help: Image: State of the and ML provide interface to help: Image: State of the and ML provide interface interface to help: Image: State of the and ML provide interface interfac
[R.opp Gul 1.65 (6313) x86_64-apple-darwin13.4.8] [Parentes Date Modified Outcoments Date Modified Outcoments Date Modified Date Created [Parentes [Parentes Parentes Parentes<
Thistory restored from //sers/stationsys/seps.history] > sources("http://wm.donielerrjohnson.com/Bbrul.R") > fnulO > fnulO With come to Brul version is 2.2.8 - High Beans Your current & version is 3.1.3 - Smoth Sidewalt In case of trubit, mode sure your & Version is subdet Please on th restruct to emil denielerrjohnsomport Disclose Offee Your current & version is 3.1.3 - Smoth Sidewalt In case of trubit, mode sure your & Version is subdet Please on th restruct to emil denielerrjohnsomport Disclose The Please sure your & version is subdet Please on th restruct to emil denielerrjohnsomport Disclose The Please sure your & version is subdet Please and the updates occur, sight-click to do this. You may are in your data: accerts, discritics, digraphs, or eve Generos Surg ("Undid multiple serving", are thy or data: accerts, discritics, digraphs, or ere your data: accerts, discritics, digraphs, or ere in your data: accerts, discritics, discrits, discrits, discrits, discritics, discritics, discritics, disc
> surcef (http://wm.danielerojohsan.cov/Binul.R*)
chrul() Reicone to Borul version 3.2.8 - High Beams Your current Kersion 3.3.1.3 - Smooth Sidemalt In case of truble, modes sure your & version is subjected Please don't kersion to as 1.1.3 - Smooth Sidemalt In case of truble, mode sure your & version is updated Please don't kersion to as 1.1.3 - Smooth Sidemalt In done a port walk werk on the side to employ for done strukture werk on the side to employ for the dots in strukture werk on the side to employ for the dots in the version side of the side to employ for the dots. In the side to employ for the dots in the version side of the side to employ for the dots. In the side tof the dots. I
Melcome to Burul version 2.28 - High Beams Your current K version is 3.1.3 - Smooth Sideralt Applications App
Weitung to write to should weis on a site of the data. The downloaded binory packages ore in NewFlag Desktop Desktop Desktop Desktop Desktop Desktop Desktop Mar 31, 2015, 1312 MM Apr 10, 2015, 1317 MM Apr 10, 2015, 1317 MM Apr 10, 2015, 1317 MM Mar 31, 2015, 1502 Mm Desktop Places don't hestite to energic like to do this regregores of the data is accents, diacritics, digraphs, or energies accent, digraphs, or energies
In case of truble, meles sure your R version is update Please don't hesitate to emil duritle cryptogram for the data. Trub or when updates occurs, discritics, digraphs, or evi- converge properly may be too complex for the data. Trub converge properly may be too complex for
Ploase don't hesitöre to emil donielærgiohnsomførtil. Dowindads DOEfraqueKitst. A pr. 6, 2015. 11:04 AM Apr. 8, 2015. 11:04 AM A
Rindows users may need to run R as administrator the right of when updates sccure. Right related to the Ray administrator the right of when updates sccure. Right relation to the right of the date. True Downloads Downloads <t< th=""></t<>
or when updates accur. Right-Click to do this. You may If you get an error soying "Invoid an Utipty string", are tin your data: accents, discritics, digraphs, or eve Remoto Disc DefrequeRSUST, 3016, 848 PM Control Line Control Line
If you get an error soying 'involid multipyte string', are in you didati accets, didatings, or eve Broul is no longer suppressing (red) convergence worthin Converge property may be too conjects for the data. True Successfully loaded teet product end with a solid convergence worthin Successfully loaded teet product end with a solid convergence worthin Successfully loaded teet product end with a solid convergence worthin Prailed_5400 ± MICROSOFTO, Prailed_5400 ± April 3, 2015, 681 PM April 3, 2015, 581 PM Mar 33, 2015, 581 PM Mar
Bruit is no longer suppressing (red) convergence working converge property may be too complex front the data. True Successfully loaded have proceeding why too complex front the data. True Successfully loaded have proceeding why too complex from the data. True Successfully loaded have proceeding why too complex from the data. True Content type 'application/x-tan' length 346922 bytes (33 demiloaded 337 K8
April 2, 2015, 736 PM Mar 31, 2015, 12:01 Converge property may be to complex for the data. True MICROSOFTO ± Tequemory.del uta April 3, 2015, 73:64 PM Mar 31, 2015, 12:01 Successfully loaded land+ pockage. trying URL Thtp://txte Prant64_5409 ± Converge property and be to complex for the data. True Mar 31, 2015, 12:01 Content type / opplication/x-tar' length 346822 bytes (33 Prant64_5409 ± Cancel Open domiloaded 337 K8 Cancel Open The downloaded binary packages are in //ar/folders/n8/5gdd2cr5/5gleyt_lwign/58008gn/T//RtmpnfMK0p/downloaded_packages Successfully loaded Multin pockage. Successfully loaded Multin pockage. Successfully loaded Multin pockage. Successfully loaded Multin pockage.
Successfully loaded text pockage. trying UBL "http://stor Content type "application/x-tar" length 346822 bytes (33 downloaded 337 K8 The downloaded binary packages are in //ar/folders/Af/Sg6256r/Sjdeyt_lengnjp88008gn/T//RtmpnfMXXp/downloaded.packages Successfully loaded Multin pockage.
Content type 'application/x-tar' iength 346922 bytes (33 dominaded 337 K8 The dowiloaded binary packages are in /kar/folders/n&jagd255Cf5jafet_LingtpB000gs/T//RtnpnfMKXp/downloaded_packages Successfull Joaded MMIn package.
domiloaded J37 K8 The domiloaded binary packages are in //ur/foldex/sAS/SpGZCRS/g4myt_lw4m/B0000gn/T//RtmpnfMMXp/domiloaded_packages Successfully loaded MuMIn package.
downloaded 337 K8 The downloaded binary packages are in /var/folders/h8/sg6d25cn5/g6myt_lw4mjp80000gn/T//RtmpnfMMXp/downloaded_packages Successfully loaded MuKin package.
The downloaded binary packages are in /var/Tolders/n&SgdaSSCnSjuHot_LHenjp88008gn/T//RtnpnfMRXp/downloaded_packages SuccessFull_binaded Mutin package.
Successfully loaded MuMIn package.
Successfully loaded lattice package.
No data loaded.
NATE MENN 1-load/sww dota 9-rest t-swit 1:1
No data loaded.
What separates the columns in the data file to open? (<-commos s-semicolons t-tabs tf-token file) Press Enter to exit, keeping current data file, if any. 1: c

Una vez seleccionado el archivo, aparecerá en la ventana de R la estructura de los datos. Es decir, verá una lista de todas las columnas y los códigos utilizados en cada columna. Este es un buen momento de revisar sus datos para asegurarse de que todo esté codificado adecuadamente. Al final de la lista, aparece el total de los tokens u ocurrencias. Asegúrese de que sea correcto este número. Si encuentra algún error (código mal escrito, códigos duplicados en una sola columna, etc.) vuelva al archivo .csv y arregle el problema, y luego, vuelva a realizar este paso.



OJO: En el menú principal de Rbrul (que aparece otra vez debajo de la lista), está a. la opción 2 que permite ajustar la codificación. Aunque es preferible arreglar la codificación en el mismo archivo .csv, también es posible utilizar esta función para hacer lo mismo. Escriba "2" y dele a "Enter". Aparecerá un segundo menú "ADJUSTING MENU" en el cual escogerá la opción "5-recode". Luego, puede escoger la columna a recodificar escribiendo el número correspondiente. Aparecen los códigos con sus números correspondientes. Simplemente, seleccione los que quiera ajustar o unir como un grupo dándole a "Enter" después de cada número. Una vez seleccionados los números deseados, dele a "Enter" otra vez y escriba el nuevo código y dele a "Enter". Puede seguir recodificando la columna escribiendo otros números o puede darle a "Enter" para completar. Cuando le da "Enter", tendrá la opción de crear una nueva columna para la nueva codificación (escriba el nombre de la nueva columna y dele a "Enter"), o puede darle a "Enter" para seguir con la columna original. Escriba "9-main menu" para volver al menú principal. Aparecerá de nuevo la lista de columnas, la cual incluirá la nueva codificación (aparece al final si es una nueva columna).



6. TABULACIONES CRUZADAS

Antes de realizar un análisis estadístico multivariado, es recomendable hacer una serie de tabulaciones cruzadas para familiarizarse con sus datos. Estas tabulaciones, también conocidas como tablas de contingencia, permiten detectar errores en la codificación y multicolinealidad (i.e., correlaciones fuertes entre variables independientes). Además, puede visualizar las distribuciones generales de la variable bajo estudio. Escriba la opción 4. Decida qué grupos factoriales o variables independientes quisiera analizar. La primera serán las columnas y la segunda, las filas o líneas. Dele a "Enter" después de cada selección. Luego, habrá una opción de "pages". Ignoraremos esta opción por el momento. Déjelo en blanco y dele a "Enter". Ahora, tendrá que decidir cómo quisiera ver los datos, como una proporción o un promedio o como una suma ("count"). Si quiere verlos como una suma, la cual es necesaria en los casos en que quiere hacer una prueba de chi cuadrado, dele a "Enter". Si quiere ver los datos como una proporción o un promedio, escriba "1". Por ahora, seleccionaremos la opción 1 puesto que las cifras tienen más significado. Dele a "Enter". Luego, tiene que escribir el número que corresponda a la columna de la variable dependiente. Dele a "Enter". Ahora, decida si la variable es binaria o continua (numérica). Este manual se trata de cómo realizar un análisis estadístico con datos numéricos así que escribirá la opción 1. Dele a "Enter". Aparecerá la tabulación cruzada. Si no la ve, corra para arriba un poco utilizando la barra a la derecha.

WIN MENU L-load/save data 2-adjust data L-crosstabs 5-modeling 6-plotting R-restore data 9-reset 0-exit

: 4
Seas-ub: factor for columns? (1-DRER 2-Filename 3-PART 4-context 5-new.word 6-word 7-variant 8-binaryvar 9-value 18-Duration.ns.. 11-COC_MDD 12-UNVOI 13-posimword 14-presegheight.bin 15-new.folseg.bin
F-folsegheight 2-fragment 33-ros 3X-ration 5-003 action 28-binaryvar 9-value 18-binaryvar 9-value

23 vss-tab: factor for rows? (1-GRUER 2-Filename 3-PART 4-context 5-new.word 6-word 7-variant 8-binaryvar 9-value 10-Duration.ms., 11-GGLMID 12-UNVDI 13-posimword 14-presegheight.bin 15-new.folseg.bin 16-segheight.bin 17-sylstress 18-wordlengtb.bin 13-free, plusi 20-log.free, plusi 21-cogstat 22-eec 24-hablante 25-wordlengtb. 26-typepreseg 27-preseg 28-typefolseg 29-folseg 30-presegheight Enter-none) -frequency 33-ros 34-START 35-END 36-DUR 37-example.folsegheight Enter-none)

24
Security Security

∴. Cross-tab: variable for cells? (1-response proportion/mean Enter-counts)

1:1 Diose response (dependent vorlable) by number (1-0RDER 2-Filename 3-PART 4-context 5-new.word 6-word 7-variant 8-binaryvar 9-value 10-Duration.ms.. 11-OOCMD 12-UNUI 13-posimword 14-presegheight.bin 15-mem.folsegh in 16-folsegheight.bin 72-vylatress 18-wordlength.bin 19-freq.plus1 28-log.freq.plus1 2-cogstat 22-eec 23-sex 24-hablante 25-wordlength 26-typepreseg 27-preseg 28-typefolseg 29-folseg 30-resegnieght 31-bissegheight 25-respect 33-avait 2-folse 7-eagn [edit 37-eagn] etc. 19-folsegheight 2-typepreseg 72-preseg 28-typefolseg 29-folseg 30-resegnieght 31-bissegheight 2-folsegheight 25-word etc. 19-folsegheight)

vpe of response? (1-continuous Enter-binary)

mean of COG MID sex ablante female male total bi 2298.844 1862.207 2080.914 mono 1569.917 1139.121 1402.198 total 1967.578 1610.808 1805.410

-

Primero que todo, en *la columna* de total, vemos que los bilingües tienen un promedio de CDG mayor que los monolingües, 2080.9 Hz vs. 1402.2 Hz. Esta observación nos indica que los bilingües reducen menos que los monolingües. Si miramos *la línea* de total, vemos que los hombres tienen un promedio de 1610.8 Hz, mientras que las mujeres tienen un promedio de 1967.6 Hz. Esto quiere decir que los hombres reducen más que las mujeres. Ahora, se puede apreciar también que los hombres monolingües son los que más reducen la /s/ con un promedio de 1139.1 Hz, luego siguen las mujeres monolingües con 1569.9 Hz, después siguen los hombres bilingües con 1862.2 Hz, y por último, las mujeres bilingües con 2298.8 Hz.

mean of COG_MID			
	sex		
hablante	female	male	total
bi	2298.844	1862.207	2080.914
mono	1569.917	1139.121	1402.198
total	1967.578	1610.808	1805.410

Se recomienda hacer tabulaciones cruzadas con todas las variables independientes. También se puede hacer tabulaciones cruzadas con las variables independientes y la variable dependiente. En las opciones de columnas y filas, escoja una variable independiente para las columnas y la variable dependiente para las filas. En este caso, necesitará una columna con la variable dependiente codificada categóricamente, por ejemplo, reducida vs. no reducida. Este tipo de tabulación permite observar si la codificación categórica corresponde a la variable continua. Por ejemplo, en la tabla abajo, se desprende que las variantes reducidas tienen un promedio de CDG menor en sílabas átonas (969.9 Hz) que las variantes no reducidas o sibilantes (5735.5 Hz). Se observa lo mismo en el caso de las sílabas tónicas. Además, las reducidas tienen un promedio menor en sílabas átonas (969.9 Hz) que en sílabas tónicas (1276.2 Hz). Por razones articulatorias, se esperan las dos observaciones.

mean of COG_MID			
	sylstress		
binaryvar	a	t	total
reduced	969.941	1276.152	1027.054
sibilant	5735.487	6181.546	5855.411
total	1671.868	2340.285	1805.410

Considero que este último ejemplo de tabulación cruzada se realiza mejor en Excel utilizando tablas Pivot.

Ahora, hagamos una tabulación cruzada con variables lingüísticas independientes y visualizando los datos como sumas. Esta vez las columnas serán de "altura del segmento siguiente" y las filas serán de "posición de /s/ en la palabra", dos predictores que han sido importantes en estudios previos. En vez de proporción, se escoge la opción de "count" dándole a "Enter".

```
MAIN MENU

1-load/Some data 2-adjust data

4-crosstabs 5-modeling 6-plotting

8-restare data 9-reset 0-exit

1:4

Cross-tab: factor for columns? (1-GRUER 2-Filename 3-PART 4-context 5-nem.word 6-word 7-variant 8-binarywar 9-value 10-Duration..ms., 11-COG_MID 12-UNVOI 13-posimord 14-presegheight.bin 15-new.folseg.bin

16-folsegheight.bin 17-sylstress 18-wordlength.bin 19-freq.plus1 20-log.freq.plus1 21-cogstat 22-eec 23-sex 24-hablante 25-wordlength 2E-typePreseg 27-preseg 28-typeFolseg 39-presegheight 31-

folsegheight.5:n 17-sylstress 18-wordlength.bin 19-freq.plus1 20-log.freq.plus1 21-cogstat 22-eec 23-sex 24-hablante 25-wordlength 2E-typePreseg 27-preseg 28-typeFolseg 39-presegheight 31-

folsegheight.bin 17-sylstress 18-wordlength.bin 19-freq.plus1 20-log.freq.plus1 21-cogstat 22-eec 23-sex 24-hablante 25-wordlength 2E-typePreseg 27-preseg 28-typeFolseg 29-folseg 39-presegheight 32-

frequency 33-ros 34-START 35-D0 36-DUR 37-example.folsegheight

1:3

Cross-tab: factor for rows?(1-ORDER 2-Filename 3-PART 4-context 5-new.word 6-word 7-variant 8-binarywar 9-value 10-Duration..ms., 11-COG_MID 12-UNVOI 13-posimord 14-presegheight.bin 15-new.folseg.bin 16-

folsegheight.bin 17-sylstress 18-wordlength.bin 19-freq.plus1 20-log.freq.plus1 21-cogstat 22-eec 23-sex 24-hablante 25-wordlength 2E-typePreseg 27-preseg 28-typeFolseg 29-folseg 30-presegheight 32-

frequency 33-ros 34-START 35-D0 36-DUR 37-example.folsegheight Enter-none)

1:

Cross-tab: for tor 10-rypestress 18-wordlength.bin 19-freq.plus1 21-cogstat 22-eec 23-sex 24-hablante 25-wordlength 2E-typePreseg 27-preseg 28-typeFolseg 39-presegheight 32-

frequency 33-ros 34-START 35-D0 36-DUR 37-example.folsegheight Enter-none)

1:

Cross-tab: variable for cells? (1-proportion/mean of COG_MID 2-change response variable Enter-counts)

1:

counts

f 266 44 75 340 310 1437

m 0 200 0 255 30 455

192
```

Lo que vemos en la tabulación producida es que hay tres casos donde la suma de variantes es 0, lo cual quiere decir que no existe, por ejemplo, una ocurrencia en el interior de la palabra seguida de cero (o pausa). Esto tiene sentido porque una /s/ dentro de palabra siempre es seguida por algo. También se observa que las ocurrencias en el interior de la palabra nunca se encuentran seguidas de vocales altas o no altas². Es decir, nunca le siguen vocales. Esto también tiene sentido porque el análisis se trata de la /s/ en posición de coda, la cual siempre es seguida de una consonante. Si le siguiera una vocal no se consideraría una /s/ en coda, y por tanto, estaría fuera del contexto variable.

Estas observaciones son hechos del sistema lingüístico en general y nos dicen que se está midiendo lo mismo cuando se considera el efecto de posición en la palabra y la altura del segmento siguiente. Una variable parece depender de la otra. Esto puede causar un problema en el análisis estadístico ya que las variables independientes deben ser, pues, independientes. Esto es muy común y hay diferentes formas de lidiar con este asunto. Una es recodificar los factores (véase Sección 5a). En este caso, podríamos clasificar los segmentos siguientes como +cor/alto vs. –cor/alto, o sea, según la posición de la lengua. Sin embargo, la categoría de "cero" presenta todavía un problema. La otra solución es simplemente dejar fuera una de las variables independientes mientras que se analice la otra. Volveremos a este asunto en otro momento.

Counts						
	folsegheight					
posinword	minus	minuscor	plus	pluscor	zero	total
f	266	446	75	340	310	1437
m	0	200	0	255	0	455
total	266	646	75	595	310	1892

Después de hacer todas las posibles tabulaciones, puede proceder al análisis estadístico multivariado.

² "minus" se refiere a vocales no altas como "a", "e" y "o". "plus" se refiere a vocales altas como "i" y "u". "minuscor" se refiere a consonantes no coronales como "p" o "k". "pluscor" se refiere a consonantes coronales como "t", "n" o "d".

7. ELABORACIÓN DE MODELOS ESTADÍSTICOS

Para diseñar el modelo estadístico, se escoge la opción 5-modeling en el menú principal. Aparecerá otro menú, "MODELING MENU". Luego, escriba "1" y dele a "Enter" para seleccionar las variables. Primero escoja la variable dependiente, en este caso, será el centro de gravedad, una variable continua. Dele a "Enter" después de escribir el número correspondiente. Después tendrá que especificar si la variable es continua o binaria. Escriba "1" para continua. Ahora tiene que escoger todas las variables independientes que quisiera incluir en el modelo. Escriba los números dándole a "Enter" después de cada uno. Al final, dele a "Enter" para seguir al próximo paso donde tendrá que especificar cuáles son las variables independientes continuas, por ejemplo, duración, frecuencia, velocidad de habla, etc. Por ahora, sólo vamos a trabajar con variables independientes categóricas, así que puede darle a "Enter" sin escribir una opción. Ahora Rbrul le pregunta si quiere incluir algunas pruebas de interacción entre las variables independientes. Ignoraremos esta parte por el momento. Dele a "Enter". Luego, le pregunta si algunas de las variables independientes son aleatorias. También ignoraremos esto por el momento. Dele a "Enter". Ahora volvemos a "MODELING MENU".

```
MUN HOW

A consists 5-nodeling 6-plotting

Servatore dots 9-reset 0-exit

1: 5

Current voriables are:

response.binory: binaryour (reduced vs. sibilant)

fixed.factor: folsepheight sylstress sex holiante

MORELING KMU

1-chocse voriables 2-one-level (recommended)

3-step-up 4-step-dom 5-step-up/step-dom 5-step-dom 5-step-up/step-dom 5-step-up/step-dom 5-step-up/step-dom 5-step-up/step-dom 5-step-up/step-dom 5-step-up/step-dom
```

Hay varias opciones de cómo podemos correr el análisis estadístico.

- a. La opción 2-one-level: un análisis en el cual se incluyen todos los factores en una corrida para determinar cuáles tienen un efecto significativo en la distribución de la variable dependiente.
- b. La opción 3-step-up: se añade un factor a la vez comenzando con el que tiene el mayor efecto en la variable dependiente. Se repite el proceso hasta que no haya más factores con un efecto significativo.
- c. La opción 4-step-down: se incluyen todos los factores en el modelo y luego Rbrul va eliminando los factores que no son significativos.
- d. La opción 5-step-up/step-down es simplemente los dos análisis anteriores en sólo paso.

Vemos en paréntesis que la opción 2-one-level es la recomendada así que solo trabajaremos con esta opción. Escriba "2" y dele a "Enter". Aparecerán los resultados del análisis.

```
MODELITING MENU
1-choose variables 2-one-level (recommended)
3-step-up 4-step-down 5-step-up/step-down
6-trim 7-plotting 8-settings 9-main menu 0-exit
10-chi-square test
1: 2
```

La primera línea de los resultados es una lista de los factores que fueron incluidos en el análisis. Después de esa línea, aparecen los grupos factoriales de nuevo con sus respectivos factores o niveles. Luego sigue información sobre el análisis en general. Vemos que hay varias cifras. Estas cifras son importantes para entender cómo está distribuida la variable según los factores.

ONE-LEVEL ANALYSIS OF RESPONSE COG_MID WITH PREDICTOR(S); folsegheight (2.15e-46) + hablante (2.41e-13) + posinword (1.46e-10) + sex (4.87e-09) + sylstress (7.8e-09) \$folsegheight
 folsegheight
 factor
 coef
 tokens
 mean

 zero
 1332.591
 310
 3125.326

 plus
 145.698
 75
 2090.127

 minus
 -203.440
 266
 1706.716

 pluscor
 -511.668
 595
 1680.277

 minuscor
 -763.181
 646
 1294.851
 \$posinword factor coef tokens m 387.593 455 2036.464 f -387.593 1437 1732.251 \$sylstress or coef tokens mean t 340.604 378 2340.285 a -340.604 1514 1671.868 \$sex factor coef tokens female 276.038 1032 1967.578 male -276.038 860 1610.808 \$hablante factor coef tokens mean bi 349.677 1124 2080.914 mono -349.677 768 1402.198 \$misc deviance AIC df intercept grand mean R2.fixed 7702110978 34184.31 9 2346.171 1805.41 0.151 Current variables are: response.continuous: COG_MID fixed.factor: folsegheight posinword sylstress sex hablante

- - Valores p: la primera línea de los resultados contiene números en paréntesis e. conocidos como valores p. Este valor indica la significancia estadística de los grupos factoriales. Un valor p por debajo de 0.05 indica que el grupo factorial condiciona significativamente la variable dependiente. En el ejemplo, se observa que todos los grupos factoriales son significativos ya que todos los valores p están por debajo de 0.05 (e-46 significa que el decimal se mueve 46 veces a las izquierda, o sea, muy por debajo de 0.05).
 - Coef (coeficientes): reflejan la fuerza de la relación entre el factor y la variable f. dependiente. Los coeficientes negativos indican una correlación negativa y los coeficientes positivos indican una correlación positiva. Mientras más grande el valor, más fuerte es la relación. Se puede expresar esa relación también en términos de favorecimiento hacia el promedio de la variable dependiente. Los coeficientes negativos indican que el factor favorece un promedio menor. Los positivos indican que el factor favorece un promedio mayor. Se observan los promedios en orden descendente para cada factor en la última columna de cada grupo. OJO: Muchos investigadores que trabajan dentro del marco variacionista

se preocupan cuando los promedios (o porcentajes en los casos de variables binarias) no coinciden con los coeficientes. Según la tradición variacionista, cuando esto ocurre, es un indicio de que hay interacción entre dos grupos factoriales. Hay varios pasos que se pueden tomar, primero, para determinar cuáles son los grupos factoriales que interactúan, y segundo, para resolver la interacción. Volveremos a este asunto en otro momento.

A modo de ejemplo, vamos a interpretar los resultados del efecto de "altura del segmento siguiente". El valor p de esta variable independiente es 2.15e-46, lo cual quiere decir que tiene un efecto significativo en la variable dependiente puesto que está por debajo de 0.05. Este hallazgo es importante para el estudio de la /s/ en coda porque se ha observado en la literatura que el contexto fonético en el que aparece el sonido afecta cómo se produce. Al considerar los coeficientes, se desprende que cuando la /s/ aparece en posición final del enunciado, o sea, hay una pausa o cero, el promedio de CDG es más alto porque el coeficiente es positivo. Es decir, este contexto fonético favorece la retención de la /s/ o un promedio más alto. Se ve también que cuando la vocal es +alta también se favorece la retención de /s/ o un promedio alto. Además, una vocal –alta favorece el debilitamiento de la /s/ o un promedio más bajo ya que el coeficiente es negativo. En el caso del punto de articulación de la consonante siguiente, se favorece un promedio más bajo tanto con +coronal como con –coronal. Note que el orden es +coronal primero y luego –coronal.

Estos hallazgos se alinean con lo que se ha visto en estudios previos y explican lo que ocurre articulatoriamente en la producción de la /s/. Con vocales altas y consonantes coronales, el ápice de la lengua se encuentra más hacia la región alveolar. Esto facilita la producción de la /s/ ya que este sonido tiene un punto de articulación alveolar. Cuando la lengua está más abajo en la boca, como en el caso de las vocales no altas y las consonantes no coronales, se dificulta la elevación de la lengua hacia los alvéolos para producir la /s/ y, por lo tanto, se debilita dicho fono.

\$folsegheight			
factor	coef	tokens	mean
zero	1332.591	310	3125.326
plus	145.698	75	2090.127
minus	-203.44	266	1706.716
pluscor	-511.668	595	1680.277
minuscor	-763.181	646	1294.851

Note que, después de la última variable independiente, hay una sección titulada \$misc. Esta es información general que provee Rbrul sobre el modelo producido. Es particularmente importante cuando diseñamos varios modelos con los mismos datos. Se utiliza esta información para comparar el ajuste de los modelos.

Deviance (desviación) – medida de cuán bien se ajusta el modelo a los datos o cuánto se desvían los datos de las predicciones del modelo. En otras palabras, mide lo bien que se

explican los datos con el modelo. Mientras más grande la desviación, peor es el ajuste. La desviación tiende a aumentar a medida de que se vayan añadiendo predictores o variables independientes al modelo.

Degrees of freedom (df – grado de libertad) – df es el número de parámetros en el modelo, una medida de la complejidad del modelo. Mientras más factores en el modelo, mayor es el grado de libertad.

Intercept (intercepto) – Es el valor estimado de la variable dependiente si x=0.

Grand mean (promedio general) – es el promedio general de todos los datos.

R2 – se refiere a la proporción de la variación explicada por el modelo.

8. COMPARACIÓN DE DIFERENTES MODELOS

Recuerde que al final de la Sección 6 hicimos una tabulación cruzada entre "altura del segmento siguiente" y "posición en la palabra" y nos percatamos de que cuando aparece la /s/ en el interior de la palabra nunca le siguen vocales ni pausas. No le siguen vocales porque siempre es una /s/ en coda, un contexto en que solo siguen consonantes; si siguiera una vocal, no sería /s/ en coda. No le siguen pausas porque está dentro de la palabra. Concluimos de esto que hay una correlación entre estas dos variables independientes, la cual se conoce como multicolinealidad. La multicolinealidad puede afectar los valores estimados o coeficientes que produce el modelo estadístico. Es decir, puede sesgar los resultados del análisis aun cuando las dos variables salen significativas. Para evitar coeficientes imprecisos, se recomienda excluir una de las variables mientras que se analiza el efecto de la otra. Esto quiere decir que tendremos dos modelos, uno con "altura del segmento siguiente" y otro con "posición en la palabra". Luego, hacemos una comparación de los dos modelos para ver cuál tiene el mejor ajuste a los datos.

Vuelva a hacer los pasos para diseñar el modelo (note que si ya se especificó la variable dependiente previamente sólo hace falta darle a "Enter"). Vamos a elaborar un modelo con todas las variables independientes previamente incluidas menos "posición en la palabra". Es decir, el modelo tendrá "altura del segmento siguiente", "acentuación silábica", "sexo" y "tipo de hablante".

```
MODELING MENU

1-choose variables 2-one-level (recommended)

3-step-up 4-step-down 5-step-up/step-down

6-trin 7-plotting 8-settings 9-main menu 0-exit

10-chi-square test
1:1 Toose response (dependent variable) by number, or Enter to keep COG_MID (1-ORDER 2-Filename 3-PART 4-context 5-new.word 6-word 7-variant 8-binaryvar 9-value 18-Duration.ms.. 11-COG_MID 12-UNNOI 13-
posimord 14-presegheight.bin 15-new.folsegh.bin 16-folsegheight.bin 17-sylstress 18-wordlength.bin 19-freq.plus1 28-tog.freq.plus2 21-cogstat 22-eee 23-sex 24-hablante 25-wordlength 26-typepreseg 27-preseg
28-typefolseg 29-folseg 38-presegheight 31-folsegheight 52-frequency 33-ros 34-START 35-END 36-DUR 37-example.folsegheight)
  1:
Choose predictors (independent variables) by number, or Enter to keep folsegheight & sylstress & sex & hablante (1-0RDER 2-Filename 3-PART 4-context 5-new.word 6-word 7-variant &-binaryvar 9-volue 10-
Durotion, ms., 12-UNVDI 13-posimord 14-presegheight 5-in 15-new.folsegheight 15-rdisguheight 5-in 15-hold 18-binaryvar 9-volue 10-
wordinght 75-typerseg 72-presega 23-typerologa 29-folseg 30-presegheight 31-folsegheight 7- von 35-TART 35-END 36-UN 72-vacemple folsegheight)
 5:
Are any predictors continuous? (31-folsegheight 17-sylstress 23-sex 24-hablante Enter-none)
 Consider a pairwise interaction of fixed effects (for interactions between fixed and random effects, use random slopes, below)? Choose two at a time.
(31-folsegheight 17-sylstress 23-sex 24-hablante Enter-done)
 Any random intercepts? (31-folsegheight 17-sylstress 23-sex 24-hablante Enter-none)
Current variables are:
response.continuous: COG_MID
fixed.factor: folsegheight sylstress sex hablante
MODELING MENU

1-choose variables 2-one-level (recommended)

3-step-up 4-step-down 5-step-up/step-down

6-trim -plotting 8-settings 9-main menu 0-exit

10-chi-square test
```

Los resultados aparecen en la imagen abajo.

ONE-LEVEL ANALYSIS OF RESPONSE COG_MID WITH PREDICTOR(S): folsegheight (8.1e-38) + hablante (3.64e-14) + sylstress (2.32e-10) + sex (4.93e-08)

\$folsegheight coef tokens factor factor coef tokens mean zero 1220.815 310 3125.326 plus 21.797 75 2090.127 pluscor -291.971 595 1680.277 minus-318.452 266 1706.716 minuscor -632.188 646 1294.851 \$sylstress factor coef tokens mean t 376.267 378 2340.285 a -376.267 1514 1671.868 \$sex factor coef tokens mean female 259.618 1032 1967.578 male -259.618 860 1610.808 \$hablante factor coef tokens mean bi 364.971 1124 2080.914 mono -364.971 768 1402.198 \$misc deviance AIC df intercept grand mean R2.fixed 7870132433 34223.14 8 2090.607 1805.41 0.133 Current variables are: response.continuous: COG_MID fixed.factor: folsegheight sylstress sex hablante

Note que todos los grupos factoriales son significativos. También note la desviación: 7870132433. Ahora vamos a elaborar otro modelo, pero esta vez, en lugar de "altura del segmento siguiente" incluiremos "posición de la palabra" junto con las otras variables independientes. Los resultados aparecen en la imagen abajo.

```
ONE-LEVEL ANALYSIS OF RESPONSE COG_MID WITH PREDICTOR(S): hablante (1.7e-13) + sylstress (1.46e-08) + sex (4.51e-06) + posinword (0.0339)

$posinword

factor coef tokens mean

m 122.343 455 2036.464

f -122.343 1437 1732.251

$sylstress

factor coef tokens mean

t 349.023 378 2340.285

a -349.023 1514 1671.868

$sex

factor coef tokens mean

female 227.629 1032 1967.578

male -227.629 1032 1967.578

mole -227.629 1032 1967.578

sisc

factor coef tokens mean

b 371.241 1124 2080.914

mono -371.241 124 2080.914

mono -371.241 768 1402.198

$misc

deviance AIC df intercept grand mean R2.fixed

8600847302 34385.12 5 1987.925 1805.41 0.052

Current variables are:

response.continuous: COG_MID

fixed.factor: posinword sylstress sex hablante
```

Vemos en los resultados del segundo modelo que todos los grupos factoriales son significativos. Note la desviación: 8600847302. Observamos que del primer modelo al segundo, aumentó la desviación, lo cual quiere decir que el segundo modelo tiene un peor ajuste. Es decir, los datos son explicados mejor con "altura del segmento siguiente", "acentuación silábica", "sexo" y "hablante" como variables independientes. Ahora queremos saber si esta diferencia entre los dos modelos es significativa. Si el primer modelo es significativamente mejor, nos quedamos con ése y descartamos el segundo. Para determinar esto, tendremos que realizar una prueba de chi cuadrado, la cual se encuentra como opción 10 en "MODELING MENU". Escriba "10" y dele a "Enter". Rbrul le pide la primera desviación. Introdúzcala y dele a "Enter". Ahora introduzca la segunda desviación y dele a "Enter". Luego Rbrul le pregunta si los valores que fueron introducidos fueron valores de "log likelihood". La respuesta es no, así que dele a "Enter". Ahora le pide que introduzca la diferencia entre los grados de libertad de los dos modelos. El df del primer modelo es 8 y el df del segundo es 5, así que la diferencia es 3. Escriba 3 y dele a "Enter". Los resultados de la prueba de chi cuadrado aparecen en la imagen abajo.

```
MODELING MENU
1-choose variables 2-one-level (recommended)
3-step-up 4-step-down 5-step-up/step-down
6-trim 7-plotting 8-settings 9-main menu 0-exit
10-chi-square test
1: 10
Enter first deviance or log likelihood.
1: 7870132433
Enter second deviance or log likelihood.
1: 8600847302
If these were log likelihood values, press 1. Press Enter if they were deviances.
1:
Enter difference in degrees of freedom.
1: 3
Chi-square = 730714869, df = 3, p = 0
```

El número que más importa aquí es el valor p. Debe ser 0.05 o menos. El valor p de la prueba es 0. Es decir, la diferencia entre los dos modelos es significativa, y el primer

modelo es significativamente mejor que el segundo. Así que descartamos el segundo modelo.

Recuerde que en la sección 7f se mencionó que muchos variacionistas se preocupan cuando los promedios no coinciden con los coeficientes. En nuestro análisis, resulta que tenemos este caso. Observe en los resultados de "altura del segmento siguiente" que "pluscor" y "minus" no están en orden descendente; "minus" debe preceder a "pluscor". Puede que haya otra variable independiente que esté interactuando con esta. Podemos determinar cuál es esa variante si hacemos una especie de "step-up" (véase 7b) manual en el que elaboramos varios modelos comenzando con "altura del segmento siguiente" (los promedios deben coincidir con los coeficientes cuando se analiza solo) y añadiendo grupos factoriales hasta que vemos el cambio en el orden. Sabremos que es el último grupo factorial añadido que interactúa con "altura del segmento siguiente". Cuando hago este procedimiento, resulta que es "acentuación silábica" la que interactúa con "altura del segmento siguiente".

```
ONE-LEVEL ANALYSIS OF RESPONSE COG_MID WITH PREDICTOR(S): folsegheight (1.82e-34)
$folseaheight
                coef tokens
   factor
                                     mean
     zero 1145.866 310 3125.326
 plus 110.667 75 2090.127
minus -272.743 266 1706.716
pluscor -299.182 595 1680.277
minuscor -684.608 646 1294.851
$misc
   deviance
                  AIC df intercept grand mean R2.fixed
 8347361567 34328.53 5 1979.459 1805.41 0.08
Current variables are:
response.continuous: COG_MID
fixed.factor: folsegheight
ONE-LEVEL ANALYSIS OF RESPONSE COG_MID WITH PREDICTOR(S): folsegheight (1.23e-35) + sylstress (6.47e-09)
$folsegheight
  factor
            coef tokens
                             mean
    zero 1201.645 310 3125.326
    plus 29.468
                      75 2090.127
                    595 1680.277
                                   pluscor -285.646
   minus -311.139 266 1706.716
minuscor -634.328 646 1294.851
$sylstress
factor coef tokens mean
t 351.128 378 2340.285
a -351.128 1514 1671.868
$misc
             AIC df intercept grand mean R2.fixed
  deviance
8200871776 34297.03 6 2168.338
                                 1805.41
                                             0.096
Current variables are:
response.continuous: COG_MID
fixed.factor: folsegheight sylstress
```

Se puede hacer una prueba de interacción para ver si es significativa o no. Cuando se elabora el modelo, Rbrul le permite escoger dos factores a la vez para hacer una prueba de interacción.



Los resultados del nuevo modelo se encuentra en la imagen abajo. Observamos que la interacción entre "acentuación silábica" y "altura del segmento siguiente" es significativa. También notamos que cuando se incluye la prueba de interacción, se arregla el orden de los promedios en "altura del segmento siguiente". A veces la interacción tiene implicaciones importantes; depende del interés del investigador. No entraremos en detalle de lo que significa la interacción en este análisis.

```
ONE-LEVEL ANALYSIS OF RESPONSE COG_MID WITH PREDICTOR(S): hablante (1.03e-14) + sex (7.6e-08) +
folsegheight:sylstress (9.33e-05) + folsegheight [main effect, not tested] + sylstress [main effect,
not tested]
$folsegheight
            coef tokens
  factor
                            mean
                   310 3125.326
    zero 1382.147
    plus 214.584
                     75 2090.127
   minus -320.772
                   266 1706.716
 pluscor -497.107
                    595 1680.277
minuscor -778.853
                   646 1294.851
$sylstress
factor
           coef tokens
                           mean
     t 564.073 378 2340.285
     a -564.073 1514 1671.868
$sex
factor
           coef tokens
                          mean
female 254.701 1032 1967.578
  male -254.701
                 860 1610.808
$hablante
           coef tokens
factor
                           mean
    bi 371.115 1124 2080.914
  mono -371.115 768 1402.198
$`folsegheight:sylstress`
factor:factor coef tokens
                                 mean
       plus:t 651.177 26 3622.931
    pluscor:a 432.775
                         469 1633.311
                       543 1218.696
   minuscor:a 309.211
      zero:t 137.115
                          47 4302.892
      minus:a 46.306
minus:t -46.306
                         190 1443.447
                         76 2364.888
      zero:a -137.115
                         263 2914.886
   minuscor:t -309.211
                         103 1696.330
    pluscor:t -432.775
                         126 1855.095
       plus:a -651.177
                          49 1276.802
$misc
   deviance
                AIC df intercept grand mean R2.fixed
7772300037 34207.48 12 2153.911 1805.41 0.143
Current variables are:
response.continuous: COG_MID
fixed.factor: folsegheight sylstress sex hablante
fixed.interaction: folsegheight:sylstress
```

9.	PRESENTACIÓN DE LOS RESULTADOS

Lo siguiente es un ejemplo de cómo se presentan los resultados en una tabla.

	Center	• of Gravity of	Coda /s/	
Total Deviance				1892 7772300037
Intercept				2153.911
	Factor	Coef	Ν	Mean
Speaker Type	9			
	Bilingual	371.115	1124	2080.914
	Monolingual	-371.115	768	1402.198
p-value	1.03e-14			
Speaker Sex				
	Female	254.701	1032	1967.578
	Male	-254.701	860	1610.808
p-value	7.6e-08			
Interaction –	Following Segment He	ight:Syllable S	Stress	
	+High:Tonic	651.177	26	3622.931
	+Coronal:Atonic	432.775	469	1633.311
	-Coronal:Atonic	309.211	543	1218.696
	Zero:Tonic	137.115	47	4302.892
	-High:Atonic	46.306	190	1443.447
	-High:Tonic	-46.306	76	2364.888
	Zero:Atonic	-137.115	263	2914.886
	-Coronal:Tonic	-309.211	103	1696.33
	+Coronal:Tonic	-432.775	126	1855.095
	+High:Atonic	-651.177	49	1276.802
p-value	9.33e-05			
Following Seg	gment Height			
	Zero/Pause	1382.147	310	3125.326
	+High	214.584	75	2090.127
	-High	-320.772	266	1706.716
	+Coronal	-497.107	595	1680.277
	-Coronal	-778.853	646	1294.851
Syllable Stres	S			
·	Tonic	564.073	378	2340.285
	Atonic	-564.073	1514	1671.868