

Regímenes latentes de error en el aprendizaje de la concordancia plural en ELE (material suplementario)

31 de agosto de 2022

1. Motivos y LDA.

Los “motivos” [*motif*] son series de símbolos discretos $S = \{S_1, \dots, S_{\mathcal{L}}\}$ pertenecientes a un alfabeto $A = \{A_1, \dots, A_n\}$, cuyo tamaño es $|A| = \mathcal{N}$. Para descubrir “motivos” se hace correr en la serie una ventana de largo \mathcal{L} . En cada ventana se normalizan los datos. Se establece un alfabeto A de tamaño $a = |A|$. Luego, se establecen B “puntos de corte” en los cuantiles de la normal sobre los datos estandarizados (restándoles la media y dividiendo por el desvío) en la ventana, tal que la serie de cortes $B = \beta_1, \dots, \beta_{a-1}$, tenga un área bajo la curva de la normal entre β_i y β_{i+1} igual a $\frac{1}{a}$. Es decir que los “puntos de corte” producen áreas de igual tamaño bajo la curva de la distribución. Los datos que caen bajo cada área definida por los cortes reciben un símbolo común del alfabeto. El procedimiento puede ser precedido por una reducción de dimensión de los n datos a w datos ($w < n$) en el cual se dividen los datos en w segmentos y se toma en cada segmento la media (Lonardi et al., [?]). Esto último no se llevó a cabo aquí. Se partió de la respuesta binaria “1” = *error*; “0” = *no error* y se la consideró como números enteros. Se estableció $A = \{a, b\}$, $|A| = 2$, $\mathcal{L} = 3$. Por ejemplo si una ventana es [011] su estandarización es $[-1,15; 0,57; 0,57]$. Hay un solo punto de corte en $p = \frac{1}{a} = 0,5$, por lo tanto el “motivo” será: “abb”. El Cuadro que sigue muestra los motivos presentes en las series de los cuatro aprendientes.

	motivo	secuencia
1	aaa	000
2	aab	001
3	aba	010
4	abb	011
5	baa	100
6	bab	101
7	bba	110
8	bbb	111

Cuadro 1: Motivos y secuencias que representan.

Para cada sesión de cada aprendiente se encontraron los “motivos”. Se consideró que cada *motivo* representaba un “estado” diferente de dinámica, con dos atractores “a = correcto” y “b = error” según: (i) atractor “a” fuerte: *aaa*; (ii) atractor “a” débil: *aab, aba, baa*; (iii) atractor “b” fuerte: *bbb*; (ii) atractor “b” débil: *bba, bab, abb*. Se interpretó “fuerte” como permanencia en el atractor y “débil” como un solo apartamiento del atractor hacia el otro. El objetivo consistió en clasificar las sesiones en grupos de dinámica similar, según “estados” similares. Para ello, se consideró a cada sesión como un “texto” cuyas palabras eran los “motivos”. Se organizaron los datos en términos de una matriz $d \times t$, con d sesiones / documentos en las filas y t “motivos” / términos en las columnas. En cada celda se contó el número de ocurrencias de un “motivo” determinado en una sesión determinada. Se asignó a cada término / “motivo” t en el “documento” / sesión d un peso según el esquema:

$$w_{t,d} = tf \times idf = tf_{t,d} \times \log_2 \left(\frac{N}{df_t} + 1 \right)$$

Donde: N es el número de documentos, df_t es el número de documentos que contienen el término t , $tf_{t,d}$ es el número de ocurrencias del término t en un documento d . No se consideró la primera sesión de SONIA por tener solamente cuatro instancias, con lo cual hubo 51 sesiones / documentos. El peso aumenta con la cantidad de ocurrencias de un “motivo” en una sesión y también con su rareza en la colección de documentos / “sesiones”. A continuación se hizo un análisis para descubrir “tópicos”, como si cada tópico fuera un grupo de dinámica diferente. Para ello se aplicó *Latent Dirichlet Allocation (LDA)* [Blei et al., [?]]. Esta técnica permite descubrir los tópicos de los que hablan los textos, asignando una probabilidad a cada tópico. En concreto, en LDA cada documento se representa como una mezcla de distribuciones sobre

tópicos latentes (no observados) y cada tópico se caracteriza por una distribución asimétrica sobre palabras, en donde solo algunas palabras de V tendrán alta probabilidad. Defínase: (i) una palabra w de un vocabulario $V = \{1, \dots, V\}$ como un vector de V componentes donde $w_i = 1$ si $V = i$ y $w_i = 0$ si $V = j$ ($i \neq j$); (ii) un documento \mathbf{w} como una secuencia de N palabras denotada por $\mathbf{w} = \{w_1, \dots, w_N\}$; (iii) un corpus D como una secuencia de M documentos denotada por $D = \{\mathbf{w}_1, \dots, \mathbf{w}_M\}$; (iv) un tópico φ de la colección $\Phi = \{\varphi_1, \dots, \varphi_K\}$. LDA es un modelo bayesiano generativo (en donde las observaciones están generadas por variables latentes), en el cual se asume que *cada* documento \mathbf{w} de largo N_i en el corpus D se genera de la siguiente manera.

1. Se eligen N palabras del documento muestreando de una distribución de *Poisson*: $N \sim \text{Pois}(\lambda)$.
2. Se elige el parámetro θ_i de la distribución del *prior* sobre los tópicos muestreando de una distribución de *Dirichlet*: $\theta_i \sim \text{Dir}(\alpha)$; $i = 1, \dots, M$.
3. Se elige el parámetro φ_k de la distribución del *prior* sobre las palabras muestreando de una distribución de *Dirichlet*: $\varphi_k \sim \text{Dir}(\beta)$; $k = 1, \dots, K$.
4. Para cada palabra j ($j \in \{1, \dots, N_i\}$) en cada documento i ($i \in \{1, \dots, M\}$):
 - elegir un tópico z_{ij} muestreando de una distribución multinomial $z_{ij} \sim M(\theta_i)$.
 - elegir una palabra w_{ij} muestreando de una distribución multinomial $w_{ij} \sim M(\varphi_{z_{ij}})$. Cada palabra w_{ij} tendrá una probabilidad condicional $p(w_{ij} | z_{ij}, \beta)$.

La probabilidad conjunta es la siguiente:

$$p(\mathbf{W}, \mathbf{Z}, \boldsymbol{\theta}, \boldsymbol{\varphi}; \alpha, \beta) = \prod_{i=1}^K p(\varphi_i; \beta) \prod_{j=1}^M p(\theta_j; \alpha) \prod_{t=1}^N p(Z_{j,t} | \theta_j) p(W_{j,t} | Z_{j,t}, \varphi)$$

donde \mathbf{W} , \mathbf{Z} , $\boldsymbol{\theta}$ y $\boldsymbol{\varphi}$ son vectores de parámetros.

Por ejemplo, si se tiene el texto: [el **lunes** desayuné con **cereales**; el **martes**, con **fruta**] y se extraen los términos en negrita (lematizando) para formar el documento: $\mathbf{w} = \{\text{lunes}, \text{cereal}, \text{martes}, \text{fruta}\}$. Se eligen entonces $N = 4$ palabras. Luego se eligen las probabilidades sobre los tópicos, por ejemplo que el documento versará un 50% sobre *tiempo* [tópico A] y 50% sobre *comida* [tópico B]. Entonces las palabras (observadas) generadas por los tópicos (variables latentes) serán: (1) “lunes” [tópico A], (2) “cereal” [tópico B], (3) “martes” [tópico A], (3) “fruta” [tópico B].

Se aplicó *LDA*¹ a la matriz $d \times t$ con pesos $tf \times idf$. Se eligieron $K = 3$ tópicos, $\alpha = 6,33$; $\beta = 0,17$. Para evaluar la agrupación encontrada, se formaron los grupos g_i [$i = 1, 2, 3$] de referencia utilizando las frecuencias relativa de error (f_{error}) de las sesiones / documentos d : $g_1 = \{d : f_{error} \in [0,21, 0,35]\}$; $g_2 = \{d : f_{error} \leq 0,2\}$; $g_3 = \{d : f_{error} > 0,35\}$. Los datos con la asignación de tópicos a cada sesión y aquellos de referencia se muestran en el apéndice del capítulo. Se calculó la tabla de confusión con los grupos predichos en las filas y aquellos de referencia en las columnas. De ella se derivaron las medidas siguientes. Si se tiene la tabla de confusión, donde: TP = “true positives”; TN = “true negatives”; FP = “false positives”; FN = “false negatives”; se pueden derivar las medidas descriptas en el Cuadro que le sigue².

	+	-
+	TP	FP
-	FN	TN

Cuadro 2: Tabla de confusión.

¹Utilizando la librería *tex2vec* de R.

²En el caso de la clase ternaria se hace una tabla de confusión de cada clase, como clase positiva, contra la suma de las otras como clase negativa.

Medidas	Fórmula	Definición
accuracy (acc)	$\frac{TP+TN}{FP+TN+TP+FN}$	proporción de clasificados correctamente del total
sensitivity (sens) [recall]	$\frac{TP}{TP+FN}$	Proporción de positivos clasificados correctamente y que lo son en el patrón de referencia
specificity (spec)	$\frac{TN}{TN+FP}$	Proporción de negativos clasificados correctamente y que lo son en el patrón de referencia
positive predictive value (ppv) [precisión]	$\frac{TP}{TP+FP}$	Proporción de positivos clasificados correctamente (TP) de todos los clasificados como positivos (TP+FP)
negative predictive value (npv)	$\frac{TN}{TN+FN}$	Proporción de negativos clasificados correctamente (TN) de todos los clasificados como negativos (TN+FN)
Kappa	$(totalAcc - randomAcc) / (1 - randomAcc)$	medida de "acuerdo" entre predichos y referencia
balance accuracy	$\frac{sens+spec}{2}$	accuracy para clases desbalanceadas
F1	$2 \times \left[\frac{precision \times recall}{precision+recall} \right]$	media armónica entre precisión y recall

Cuadro 3: Medidas derivadas de la tabla de confusión.

2. Hidden Markov Models (HMM) y simulación.

Un HMM simple de primer orden se resume con las ecuaciones que siguen (Bolano, 2017):

$$a_{ij} = p(S_t = i \mid S_{t-1} = j); t = 1, \dots, T$$

$$\pi_i = p(S_{t=1} = i); i = 1, \dots, k$$

$$p_i(y_t) = p(Y_t = y_t \mid S_t = i); i = 1, \dots, k$$

Las primeras dos ecuaciones representan la parte no observable del modelo y la última, la observable. La primera ecuación indica el componente a_{ij} de la matriz de transición $A_{k \times k}$; y dice que la probabilidad del estado oculto i en el tiempo t depende del estado j en el tiempo anterior. Es decir que la variable latente S_t sigue un proceso de Markov de primer orden ya que el estado oculto en el tiempo t depende del estado previo en $t - 1$. La segunda ecuación indica la probabilidad del estado oculto i -ésimo en el primer tiempo $t = 1$. La tercera ecuación da las probabilidades de emisión: la probabilidad de que la variable aleatoria respuesta Y_t se realice en el nivel y_t dado que el proceso se encuentra en el estado oculto i -ésimo. Notar que las observaciones se consideran condicionalmente independientes dado el proceso latente porque no dependen de observaciones o estados ocultos anteriores (las respuestas son dependientes entre sí pero son independientes dado el estado oculto).

Los resultados de la simulación se evaluaron según la divergencia de Kullback-Leibler. Para distribuciones de probabilidad P y Q de una variable aleatoria discreta su divergencia es:

$$D_{KL}(P \mid Q) = \sum_i P(i) \log_2 \left(\frac{P(i)}{Q(i)} \right)$$

La divergencia KL solamente se define si P y Q suman 1 y si $Q(i) > 0$ para cualquier i tal que $P(i) > 0$. La cantidad $0 \log_2(0)$ se interpreta como cero.

3. Resultados (tablas no mostradas en el artículo)

3.1. LDA: Tópicos predichos y grupos de error.

El Cuadro 4 muestra: las sesiones, la frecuencia relativa de error en cada sesión («freq.rel»), la frecuencia absoluta de error en cada sesión («freq.abs»), la cantidad total de errores por sesión («n»), los tópicos asignados por LDA y los grupos definidos según los conjuntos de error («ref»).

3.2. HMM: Probabilidades de emisión.

Los Cuadros 5 a 8 muestran las probabilidades de emisión ilustradas en el artículo mediante grafos. Las filas contienen los estados observados («motivos») y las columnas, los estados ocultos («E1» es «estado oculto 1» e ídem con los demás). El cero indica probabilidad muy baja ($p < 0,0001$). Deben leerse por columna. Por ejemplo en SONIA, si la dinámica se encuentra en el estado oculto 2 «E2» (asociado al atractor «correcto»), entonces: la probabilidad de ver el estado observado

«000» es 0.95; la probabilidad de ver el estado observado «100» es 0.0197 y la probabilidad de ver el estado observado «001» es 0.0303. Por ende con alta probabilidad el estado que una simulación recrearía sería el «000».

3.3. HMM: Betas estimados.

Los Cuadros 9 a 20 indican los betas estimados del modelo de las regresiones multinomiales:

$$\log \left(\frac{a_{ij}}{a_{i1}} \right) = \eta_j; a_{ij}(t) = P(S_t = j | S_{t-1} = i); j = 2, 3$$

$$\eta_j = \beta_0 + \beta_1 LDA + \beta_2 EP_1 + \beta_3 EP_2 + \beta_4 GRAM + \beta_5 MOD_1 + \beta_6 MOD_2 + \beta_7 MOD_3 + \beta_8 IMA.CONC + \beta_9 FAM.LEX + \beta_{10} AN + \beta_{11} EST1 + \beta_{12} EST2 + \beta_{13} EST5 + \beta_{14} FREQ.S$$

en donde a_{ij} es una entrada de la matriz de probabilidades de transición desde el estado oculto i al estado oculto j del HMM; η_j es el predictor lineal correspondiente a la transición. La categoría de base del modelo es el estado $j = 1$. Por ejemplo, en SONIA, el siguiente es el modelo desde el estado oculto $i = 1$ al estado oculto $j = 2$ es (ver segunda columna del Cuadro 9; no hay datos para $MOD = \langle 1 \rangle$):

$$\log \left(\frac{a_{12}}{a_{11}} \right) = \eta_2 = -118,2289 + 65,6168LDA + 42,1935EP_1 - 12,6899EP_2 - 66,4933GRAM + \beta_5(MOD_1 = 0) + 86,1733MOD_2 + 41,9845MOD_3 + 44,0139IMA.CONC - 64,9806FAM.LEX - 22,9882AN + 152,9553EST1 + 86,1287EST2 + 93,9047EST5 + -131,8696FREQ.S$$

Por ende el logit de la probabilidad de pasar del estado oculto uno al estado oculto dos sube con un beta estimado de signo positivo y baja con uno negativo. Entonces contribuyen a que el modelo abandone el estado oculto uno las variables: LDA, EP = $\langle 1 \rangle$, MOD = $\langle 2 \rangle$, $\langle 3 \rangle$, IMA.CONC, EST1, EST2, EST5. En el caso de que todas las predictoras valieran cero dicho modelo sería simplemente: $\log \left(\frac{a_{12}}{a_{11}} \right) = \eta_2 = \beta_0 = -118,2289$, en donde el logit de la probabilidad de pasar del estado oculto uno al estado oculto dos disminuye 118 unidades. Las probabilidades de transición mostradas en el artículo se calculan como:

$$P(S_t = j | S_{t-1} = i) = \frac{e^{\eta_j}}{\sum_{j=1}^3 e^{\eta_j}}; j = 1, 2, 3$$

Por ejemplo en el caso de que todas las predictoras valieran cero, la probabilidad de permanecer en el estado 2 (asociado al atractor «correcto») sería relativamente alta [ver Cuadro 10]:

$$P(S_t = 2 | S_{t-1} = 2) = \frac{e^{\beta_{0,22}}}{\sum_{j=1}^3 e^{\beta_{0,2j}}} = \frac{e^{1,2312}}{e^0 + e^{1,2312} + e^{-62,7528}} \approx 0,77$$

En cambio, si ahora se agrega el efecto parcial de una concordancia a larga distancia (LDA = $\langle 1 \rangle$), dejando las demás predictoras en cero, entonces la misma probabilidad sería [ver Cuadro 10]:

$$\begin{aligned} \eta_{21} &= \beta_0 + \beta_1(LDA = 1) = \beta_0 + \beta_1 = 0 + 0 = 0 \\ \eta_{22} &= \beta_0 + \beta_1(LDA = 1) = \beta_0 + \beta_1 = 1,2312 + 9,6669 = 10,8981 \\ \eta_{23} &= \beta_0 + \beta_1(LDA = 1) = \beta_0 + \beta_1 = -62,7528 + 103,6410 = 40,8882 \end{aligned}$$

$$P(S_t = 2 | S_{t-1} = 2) = \frac{e^{\eta_2}}{\sum_{j=1}^3 e^{\eta_j}} = \frac{e^{10,8981}}{e^0 + e^{10,8981} + e^{40,8882}} \approx 9,45 \times 10^{-14} < 0,0001$$

Con lo cual una concordancia a larga distancia hace que la dinámica deje el estado oculto asociado al atractor «correcto», ejerciendo una influencia negativa.

3.4. Simulación

La tabla 21 muestra las probabilidades empíricas y simuladas de cada estado observado, según cada sujeto. Esta tabla fue graficada en el artículo.

	sesiones	freq.rel	freq.abs	n	temas	ref.
4	1.2	0.33	6.00	18	2	2.00
5	1.3	0.21	4.00	19	1	2.00
6	1.4	0.33	14.00	42	2	2.00
7	1.5	0.26	11.00	42	1	2.00
8	1.6	0.28	7.00	25	2	2.00
9	1.7	0.11	3.00	28	1	1.00
10	1.8	0.19	7.00	37	1	1.00
11	1.9	0.19	5.00	26	1	1.00
1	1.10	0.05	1.00	19	1	1.00
2	1.11	0.18	6.00	34	2	1.00
3	1.12	0.06	1.00	18	1	1.00
12	2.1	0.40	12.00	30	3	3.00
18	2.2	0.62	10.00	16	3	3.00
19	2.3	0.24	5.00	21	2	2.00
20	2.4	0.15	2.00	13	1	1.00
21	2.5	0.33	6.00	18	2	2.00
22	2.6	0.38	18.00	47	3	3.00
23	2.7	0.35	13.00	37	2	3.00
24	2.8	0.50	19.00	38	3	3.00
25	2.9	0.26	7.00	27	2	2.00
13	2.10	0.38	14.00	37	3	3.00
14	2.11	0.24	7.00	29	2	2.00
15	2.12	0.32	9.00	28	2	2.00
16	2.13	0.27	6.00	22	2	2.00
17	2.14	0.11	3.00	27	1	1.00
26	3.1	0.23	6.00	26	1	2.00
32	3.2	0.19	10.00	52	1	1.00
33	3.3	0.03	1.00	30	1	1.00
34	3.4	0.21	6.00	28	2	2.00
35	3.5	0.20	8.00	41	1	1.00
36	3.6	0.07	2.00	30	1	1.00
37	3.7	0.24	9.00	37	2	2.00
38	3.8	0.25	8.00	32	2	2.00
39	3.9	0.16	8.00	51	1	1.00
27	3.10	0.28	5.00	18	2	2.00
28	3.11	0.15	7.00	47	1	1.00
29	3.12	0.03	1.00	32	1	1.00
30	3.13	0.19	5.00	27	1	1.00
31	3.14	0.00	0.00	13	1	1.00
40	4.1	0.15	12.00	80	1	1.00
44	4.2	0.33	13.00	39	2	2.00
45	4.3	0.33	21.00	64	2	2.00
46	4.4	0.24	15.00	62	1	2.00
47	4.5	0.24	8.00	34	2	2.00
48	4.6	0.21	10.00	47	2	2.00
49	4.7	0.37	30.00	81	3	3.00
50	4.8	0.20	7.00	35	1	1.00
51	4.9	0.36	27.00	74	3	3.00
41	4.10	0.33	17.00	52	2	2.00
42	4.11	0.46	22.00	48	3	3.00
43	4.12	0.25	19.00	75	2	2.00

Cuadro 4: Clasificación de tópicos. sesiones: sesión por cada alumno; freq.rel: frecuencia relativa de error; freq. abs: frecuencia absoluta de error; n: número de instancias por sesión; tópicos: tópico asignado a cada sesión; ref.: tópico de referencia

	E1	E2	E3
000	0.0000	0.9500	0.0169
100	0.2166	0.0197	0.2185
010	0.2235	0.0000	0.3108
001	0.3705	0.0303	0.0000
111	0.0000	0.0000	0.0672
110	0.0000	0.0000	0.2017
011	0.0000	0.0000	0.1849
101	0.1895	0.0000	0.0000

Cuadro 5: SONIA : Probabilidades de transición

	E1	E2	E3
000	0.7258	0.0000	0.0000
100	0.0000	0.5439	0.0000
010	0.0000	0.0000	0.3576
001	0.2382	0.0000	0.0000
111	0.0000	0.2363	0.0692
110	0.0000	0.0000	0.3260
011	0.0360	0.0000	0.2472
101	0.0000	0.2199	0.0000

Cuadro 6: NATI : Probabilidades de transición

	E1	E2	E3
000	0.0000	1.0000	0.0000
100	0.0000	0.0000	0.4442
010	0.5498	0.0000	0.0000
001	0.0000	0.0000	0.4442
111	0.0282	0.0000	0.0000
110	0.2115	0.0000	0.0000
011	0.2106	0.0000	0.0107
101	0.0000	0.0000	0.1009

Cuadro 7: JAKO : Probabilidades de transición

	E1	E2	E3
000	0.0000	0.0000	0.7148
100	0.0000	0.0000	0.2321
010	0.4484	0.0000	0.0000
001	0.0000	0.6618	0.0000
111	0.0897	0.1324	0.0000
110	0.1608	0.0000	0.0531
011	0.3011	0.0000	0.0000
101	0.0000	0.2059	0.0000

Cuadro 8: MIRKA : Probabilidades de transición

	St1	St2	St3
(Intercept)	0.0000	-118.2289	9.3223
LDA1	0.0000	65.6168	131.9434
EP1	0.0000	42.1935	-65.6624
EP2	0.0000	-12.6899	-19.2614
GRAM1	0.0000	-66.4933	98.5146
MOD2	0.0000	86.1733	15.4728
MOD3	0.0000	41.9845	-73.7050
IMA_CONC1	0.0000	44.0139	-52.2018
FAM_LEX1	0.0000	-64.9806	84.1449
AN1	0.0000	-22.9882	17.4624
EST11	0.0000	152.9553	-148.8821
EST21	0.0000	86.1287	-199.7842
EST51	0.0000	93.9047	-84.1072
FREQ_S1	0.0000	-131.8696	116.2469

Cuadro 9: SONIA : Betas estimados (modelo desde estado 1)

	St1	St2	St3
(Intercept)	0.0000	1.2312	-62.7528
LDA1	0.0000	9.6669	103.6410
EP1	0.0000	3.5322	53.1631
EP2	0.0000	-16.4116	-11.3476
GRAM1	0.0000	2.4153	11.4159
MOD2	0.0000	-1.7511	-16.5701
MOD3	0.0000	-1.7084	-43.5297
IMA_CONC1	0.0000	-0.9602	-16.1932
FAM_LEX1	0.0000	-0.6606	-66.6037
AN1	0.0000	3.2413	50.8271
EST11	0.0000	1.1687	-11.8636
EST21	0.0000	3.1684	21.9424
EST51	0.0000	10.0111	-5.8087
FREQ_S1	0.0000	-0.0452	20.5203

Cuadro 10: SONIA : Betas estimados (modelo desde estado 2)

	St1	St2	St3
(Intercept)	0.0000	-37.2101	-78.8184
LDA1	0.0000	-4.1986	117.3293
EP1	0.0000	73.8786	4.6306
EP2	0.0000	0.8352	-147.2660
GRAM1	0.0000	10.4430	105.1990
MOD2	0.0000	7.2356	69.5219
MOD3	0.0000	-44.0582	97.0738
IMA_CONC1	0.0000	8.4174	-70.0388
FAM_LEX1	0.0000	16.6606	-36.1525
AN1	0.0000	-103.2983	1.2940
EST11	0.0000	96.7609	-9.0716
EST21	0.0000	63.0093	-89.1719
EST51	0.0000	-3.9539	-40.1337
FREQ_S1	0.0000	-50.3814	87.5996

Cuadro 11: SONIA : Betas estimados (modelo desde estado 3)

	St1	St2	St3
(Intercept)	0.0000	-83.5512	-0.6191
LDA1	0.0000	1.3000	-0.8444
EP1	0.0000	28.5363	-0.1855
EP2	0.0000	10.0841	-6.0689
GRAM1	0.0000	32.3133	0.1064
MOD1	0.0000	-1.8385	16.2258
MOD2	0.0000	-13.3040	0.9308
MOD3	0.0000	-9.6695	1.3185
IMA_CONC1	0.0000	-5.4750	-0.0317
FAM_LEX1	0.0000	10.1678	-0.9521
AN1	0.0000	3.4661	0.4609
EST11	0.0000	-10.3355	-1.8809
EST21	0.0000	35.5469	-1.2031
EST51	0.0000	-10.4104	-1.4614
FREQ_S1	0.0000	37.9854	0.9461

Cuadro 12: NATI : Betas estimados (modelo desde estado 1)

	St1	St2	St3
(Intercept)	0.0000	-33.3340	1.6412
LDA1	0.0000	-9.1511	-0.9398
EP1	0.0000	-4.6371	1.5392
EP2	0.0000	-9.7291	-11.8739
GRAM1	0.0000	-50.2823	-5.2281
MOD1	0.0000	-1.7953	9.9183
MOD2	0.0000	23.5679	2.7772
MOD3	0.0000	-21.0371	-0.5149
IMA_CONC1	0.0000	24.0748	-0.7917
FAM_LEX1	0.0000	-20.9929	0.1586
AN1	0.0000	20.9117	1.4409
EST11	0.0000	-22.2439	-1.0729
EST21	0.0000	24.9262	0.5121
EST51	0.0000	-2.8246	-1.4095
FREQ_S1	0.0000	-26.3795	-2.8692

Cuadro 13: NATI : Betas estimados (modelo desde estado 2)

	St1	St2	St3
(Intercept)	0.0000	27.5975	22.6675
LDA1	0.0000	-2.2388	11.0442
EP1	0.0000	-7.5468	-0.2323
EP2	0.0000	5.9552	-4.7058
GRAM1	0.0000	11.3410	6.6243
MOD1	0.0000	-2.1477	8.4168
MOD2	0.0000	0.9506	-0.3737
MOD3	0.0000	7.9998	5.4493
IMA_CONC1	0.0000	21.6639	16.0351
FAM_LEX1	0.0000	-17.4035	-15.4697
AN1	0.0000	-19.8592	-12.4840
EST11	0.0000	-4.5058	-1.4142
EST21	0.0000	-8.8188	-4.2018
EST51	0.0000	-1.3959	-14.8465
FREQ_S1	0.0000	10.9507	7.7611

Cuadro 14: NATI : Betas estimados (modelo desde estado 3)

	St1	St2	St3
(Intercept)	0.0000	-60.2196	3.6808
LDA1	0.0000	16.3231	-5.3122
EP1	0.0000	-12.0976	14.8843
EP2	0.0000	0.8689	11.2981
GRAM1	0.0000	6.9741	-1.4486
MOD1	0.0000	25.1394	5.3182
MOD2	0.0000	-17.0259	-2.7087
MOD3	0.0000	13.1326	-2.7275
IMA_CONC1	0.0000	39.9705	0.1792
FAM_LEX1	0.0000	40.8157	-0.0880
AN1	0.0000	-32.1799	-0.9303
EST11	0.0000	6.7952	0.8186
EST21	0.0000	-19.2500	-0.5618
EST51	0.0000	-10.4858	-14.3697
FREQ_S1	0.0000	9.4530	1.5081

Cuadro 15: JAKO : Betas estimados (modelo desde estado 1)

	St1	St2	St3
(Intercept)	0.0000	21.2783	19.4282
LDA1	0.0000	-0.1960	-10.8494
EP1	0.0000	-6.3903	-6.4807
EP2	0.0000	1.8796	-9.1092
GRAM1	0.0000	-7.5743	-7.3945
MOD1	0.0000	-0.9927	-0.0451
MOD2	0.0000	-2.5618	-2.7484
MOD3	0.0000	8.3344	8.4542
IMA_CONC1	0.0000	-0.1216	-0.9052
FAM_LEX1	0.0000	7.6210	6.7644
AN1	0.0000	2.9786	3.7297
EST11	0.0000	-0.1164	0.1743
EST21	0.0000	1.2188	2.1258
EST51	0.0000	5.5170	5.7189
FREQ_S1	0.0000	5.7548	5.6784

Cuadro 16: JAKO : Betas estimados (modelo desde estado 2)

	St1	St2	St3
(Intercept)	0.0000	-0.4036	-2.5992
LDA1	0.0000	-8.9441	-10.3568
EP1	0.0000	0.1869	-2.1606
EP2	0.0000	0.4346	-1.0709
GRAM1	0.0000	0.3048	-3.5920
MOD1	0.0000	2.5347	13.3114
MOD2	0.0000	-0.7495	-12.9755
MOD3	0.0000	-0.5749	0.8545
IMA_CONC1	0.0000	-0.8037	3.2075
FAM_LEX1	0.0000	0.9755	0.2402
AN1	0.0000	-0.0712	0.6734
EST11	0.0000	-0.2300	1.0479
EST21	0.0000	-0.2981	-11.7190
EST51	0.0000	0.2330	2.0854
FREQ_S1	0.0000	0.3677	0.0648

Cuadro 17: JAKO : Betas estimados (modelo desde estado 3)

	St1	St2	St3
(Intercept)	0.0000	59.1342	60.1009
LDA1	0.0000	70.3073	70.2814
EP1	0.0000	29.0124	27.0848
EP2	0.0000	10.3264	8.9112
GRAM1	0.0000	90.2131	91.0007
MOD1	0.0000	1.5604	3.9374
MOD2	0.0000	-105.6143	-103.9638
MOD3	0.0000	-67.9051	-67.8772
IMA_CONC1	0.0000	29.7958	30.1893
FAM_LEX1	0.0000	-43.2763	-43.8203
AN1	0.0000	-1.7380	-1.9623
EST11	0.0000	-38.3257	-39.6199
EST21	0.0000	72.0789	70.1137
EST51	0.0000	9.7845	11.3127
FREQ_S1	0.0000	58.9295	60.4228

Cuadro 18: MIRKA : Betas estimados (modelo desde estado 1)

	St1	St2	St3
(Intercept)	0.0000	-22.2451	-59.2738
LDA1	0.0000	-6.5145	59.7950
EP1	0.0000	-7.4985	-21.6084
EP2	0.0000	-4.1354	99.8636
GRAM1	0.0000	-17.7908	-13.2262
MOD1	0.0000	5.3856	7.1706
MOD2	0.0000	-13.9679	-18.7407
MOD3	0.0000	-14.5659	-63.7427
IMA_CONC1	0.0000	-7.3594	27.0602
FAM_LEX1	0.0000	11.3146	28.6848
AN1	0.0000	-18.0450	19.8601
EST11	0.0000	-5.8612	2.4092
EST21	0.0000	11.9663	-19.7591
EST51	0.0000	22.8246	28.9729
FREQ_S1	0.0000	-1.0360	-9.6559

Cuadro 19: MIRKA : Betas estimados (modelo desde estado 2)

	St1	St2	St3
(Intercept)	0.0000	48.5459	49.0584
LDA1	0.0000	-2.3273	-1.4492
EP1	0.0000	9.0070	9.5379
EP2	0.0000	-7.4541	6.3509
GRAM1	0.0000	10.0991	10.6232
MOD1	0.0000	-5.5830	-3.3792
MOD2	0.0000	-19.9941	-18.8588
MOD3	0.0000	-8.0330	-8.1439
IMA_CONC1	0.0000	-11.5073	-11.5022
FAM_LEX1	0.0000	7.5203	7.5334
AN1	0.0000	-13.9671	-13.7451
EST11	0.0000	13.7517	13.5271
EST21	0.0000	-3.8217	-4.3662
EST51	0.0000	2.0694	0.8901
FREQ_S1	0.0000	-17.0806	-16.3954

Cuadro 20: MIRKA : Betas estimados (modelo desde estado 3)

	000	001	010	011	100	101	110	111
datos SONIA	0.510	0.112	0.122	0.038	0.112	0.049	0.042	0.014
sim SONIA	0.406	0.133	0.154	0.042	0.150	0.052	0.042	0.021
datos NATI	0.362	0.119	0.094	0.083	0.130	0.052	0.086	0.075
sim NATI	0.279	0.094	0.124	0.108	0.169	0.055	0.088	0.083
datos JAKO	0.610	0.101	0.089	0.037	0.101	0.023	0.034	0.005
sim JAKO	0.445	0.124	0.110	0.080	0.158	0.030	0.046	0.007
datos MIRKA	0.402	0.135	0.105	0.070	0.130	0.042	0.068	0.048
sim MIRKA	0.334	0.114	0.168	0.100	0.112	0.043	0.082	0.045

Cuadro 21: Probabilidades empiricas y simuladas de cada estado observado.