

# ANÁLISIS FACTORIAL DE CORRESPONDENCIAS

## Aplicación al vocabulario de una muestra de prensa (1939-45)

LYDIA ROMEU

Centro de Lexicometría  
Universidad de Barcelona

### INTRODUCCIÓN

Este artículo presenta una primera aproximación al análisis de los discursos de *Ya* y *Arriba* a partir de la descripción cuantitativa, exhaustiva, del vocabulario de una muestra de prensa correspondiente a los primeros años del franquismo.

Trataremos de reconocer tendencias globales en la evolución comparada de los discursos de la prensa falangista y católica mediante procedimientos contrastivos automatizados basados en los principios de no-lematización y segmentación del texto en formas gráficas, sistemas informáticos desarrollados por la URL «Lexicométrie et textes politiques» de la ENS de Fontenay-Saint-Cloud.

#### 1. PROBLEMÁTICA Y CRITERIOS DE MUESTREO

El *corpus* destinado a los análisis cuantitativos está constituido por una muestra de editoriales destinada a aportar datos sobre los fenómenos de frecuencia y distribución de formas presentes en los textos. La problemática general que ha guiado la constitución del *corpus* es la inscripción de los aspectos ideológicos en la discursividad. Por ello, hemos formulado la hi-

pótesis siguiente de carácter histórico: los textos de *Ya* y *Arriba*, en tanto que discursos de redistribución y reformulación de otros discursos presentarían una notable homogeneidad debida al dirigismo del Estado en materia de prensa (sistema de consignas, censura, etc.). Los cambios ideológicos (del nacionalsindicalismo al nacionalcatolicismo) que intervienen durante el período de referencia (1939-1945) permitirían, sin embargo, esperar cambios importantes en el aspecto diacrónico.

Los criterios de muestreo corresponden por lo tanto al deseo de estudiar la evolución del vocabulario y también la semejanza y disimilitud entre ambos diarios. Para ello, hemos determinado un cierto número de variables y de constantes. Entre estas, el género editorial y la homogeneidad temática, entre las variables, acontecimientos y fechas. Las tomas se han realizado de forma regular a lo largo de estos siete años.

La pequeña muestra cuenta con 36.949 ocurrencias. Es un *corpus* equilibrado desde el punto de vista del volumen tanto por lo que respecta al parámetro tiempo (tomas anuales comprendidas entre 2.128 y 3.205 ocurrencias) como a los emisores (17.089 ocurrencias de *Ya* y 19.675 de *Arriba*).

## 2. ANÁLISIS LEXICOMÉTRICOS

La lexicometría reúne una serie de métodos<sup>1</sup> que permiten, a partir de la segmentación del texto en formas gráficas, describir cuantitativamente el vocabulario de las distintas partes de un *corpus* y establecer comparaciones entre ellas. Para ello se procede, en una primera fase a la reorganización automática de la secuencia textual. Los primeros resultados proporcionan índices exhaustivos como son los índices de frecuencias, concordancias, etc. Se trata, posteriormente, de sintetizar esta información en términos contrastivos mediante análisis estadísticos, lo cual es posible gracias a la gran capacidad de cálculo del ordenador. Los métodos multidimensionales proceden a la jerarquización de los fenómenos de frecuencia a partir de índices estadísticos y a selecciones de acuerdo con umbrales preestablecidos.

---

<sup>1</sup> Sobre los métodos lexicométricos, ver A. SALEM: *Pratique des segments répétés. Essais de statistique textuelle*. Paris, Klincksieck, 1987, y L. LEBART, A. SALEM: *Analyse statistique de données textuelles*. Paris, Bordas, 1988.

### 3. EL ANÁLISIS FACTORIAL DE CORRESPONDENCIAS

Para examinar cómo se comportan las dos series cronológicas formadas por editoriales de los diarios de referencia, confrontadas a las variaciones de vocabulario, y tratar de reconocer en ellas tendencias globales, hemos sometido el *corpus* al análisis factorial de correspondencias (AFC), método estadístico que establece una tipología en términos de parentesco y disimilitud de las partes de un *corpus*, de acuerdo con el vocabulario empleado y su frecuencia.

El análisis parte del cuadro léxico que reúne los datos relativos a la repartición de las formas gráficas en las diferentes partes del *corpus*. El cuadro 1 reproduce un fragmento relativo a las 28 formas más frecuentes del *corpus* que comprende 28 líneas y 14 columnas (*Ya* y *Arriba* durante 7 años).

El objeto del análisis factorial consiste en substituir un cuadro como este índice, de doble entrada (líneas de formas gráficas y columnas de emisores) difícilmente comprensible, dada la masa de datos que contiene, por uno de fácil lectura y aportar, al mismo tiempo, una representación gráfica de los datos sintetizados, rápidamente visualizables, perdiendo en todo ello el mínimo de información posible.

El método consiste en calcular, a partir del cuadro léxico, la media de empleo de las formas del *corpus*, es decir, los «efectivos teóricos». Esta norma, confrontada a los datos realmente verificados en el *corpus* produce un nuevo cuadro de distancias, la diferencia con respecto a lo que se llama la «situación de independencia» (situación ideal que, en nuestro caso, supondría una repartición uniforme de las formas empleadas por ambos emisores en el transcurso del tiempo). El procedimiento factorial consiste en someter el cuadro resultante de las detracciones a un algoritmo, para extraer el «primer factor» que constituye la aproximación más fiel del cuadro original. El interés de la descomposición del cuadro original reside en hacer posible una descripción de las asociaciones estadísticas entre elementos línea y elementos columna y, además proporcionar una documentación que traduce gráfica y numéricamente las similitudes entre elementos, acompañando todo ello con indicadores sobre la calidad de la aproximación.

No obstante, este primer factor no es más que una aproximación a los datos reales. La diferencia entre éste y la realidad será sometida a una nueva extracción factorial. Este nuevo factor constituye una corrección del primero en el que se tendrán en cuenta los errores cometidos. Y así sucesivamente. Los factores van paulatinamente decreciendo en interés, en la medida en que no son siempre más que correcciones de los anteriores.

Los resultados relacionan por lo tanto las dos dimensiones del cuadro y señalan qué vocabulario tiende a ser utilizado por los mismos emisores



en las mismas proporciones por una parte y, por otra, qué emisores tienden a emplear el mismo vocabulario, igualmente en términos de frecuencia. El AFC es, por tanto, un instrumento especialmente adecuado para tratar de aproximarnos a las variaciones del vocabulario del *corpus*.

Presentaremos sucesivamente los resultados obtenidos y trataremos de darles una significación discursiva.

### 3.1. El escalograma de dimensiones $14 \times 2.659$

En colaboración con A. Salem, del Laboratorio de Lexicometría y textos políticos de la ENS de Fontenay-Saint-Cloud, hemos sometido el *corpus* a una primera experiencia. Para ellos hemos reagrupado los editoriales de cada locutor por año, lo cual dividió el *corpus* en 14 emisores-año (*Ya y Arriba*, de 1939 a 1945). Las formas retenidas fueron todas las formas repetidas al menos una vez en el *corpus* (únicamente fueron desechados los hapax o formas de frecuencia 1). El cuadro léxico sometido al análisis cuenta por lo tanto con 14 columnas y 2.659 líneas.

#### 3.1.1. *Emisores*

Los resultados aportan información, como decíamos, sobre el parentesco y las disimilitudes de las partes. El gráfico 1 reproduce la representación gráfica proporcionada por el ordenador relativa a los 14 emisores<sup>2</sup>. Esta representación permite reagrupar los dos primeros factores los cuales son fundamentales, dada la importancia decreciente de los mismos. Reúnen, en nuestro caso, el 20 por 100 de la información.

La tipología que construye el análisis tiende a reagrupar las partes del *corpus* que emplean el mismo vocabulario en las mismas proporciones. Únicamente mencionaremos aquí las dos constataciones principales que ha permitido el estudio de los gráficos y de los indicadores numéricos procedentes del análisis<sup>3</sup>. En primer lugar, la proximidad de los diarios del mismo año. En segundo lugar, que los puntos próximos en el tiempo están, en general,

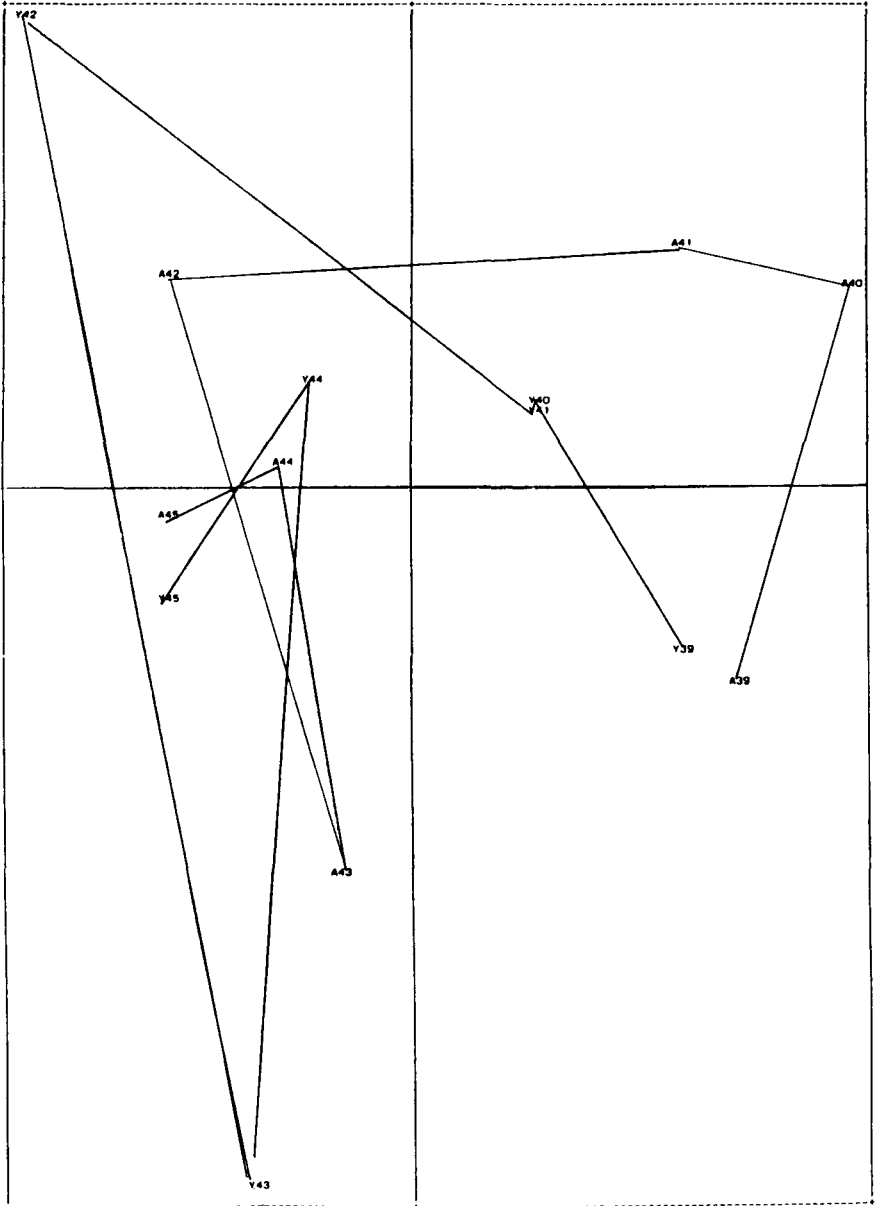
---

<sup>2</sup> En el plano la proximidad entre puntos debe interpretarse como similitudes tanto en cuanto a las formas empleadas como a su frecuencia. En el gráfico 1 hemos unido los puntos emisores-año por una línea que describe su evolución.

<sup>3</sup> Este aspecto se encuentra desarrollado en L. ROMEU: «Análisis del vocabulario de editoriales de prensa» in «Análisis estadístico de datos textuales». *Actas de V congreso de lenguajes naturales y lenguajes formales*. Universidad de Barcelona, en prensa.

# GRÁFICO 1

V 1- 0.1415 10.14% XX- 1.2991 CORPUS PRESSE FRANQUISTE ( 14 JOURNEAUX X ANNEES ) 2659 FORMES REPETÉES  
V 2- 0.1287 8.00% YY- 1.5207 CORRESPONDENCE ( 1° 2 03042100 14 POINTS ) PLOT# 1



más próximos en el plano que aquellos que están más alejados, es decir, que corresponden globalmente a lo que A. Salem llama el esquema del «tiempo léxico»<sup>4</sup>, es decir que los vocabularios de las partes correspondientes a períodos consecutivos presentan una mayor similitud que los vocabularios más alejados en el tiempo. Ello se debe a una renovación gradual del vocabulario. La curva descrita señala, sin embargo, que esta renovación no se produce de una forma gradual y ordenada. Intervienen irregularidades en 1942 y en 1943 que interrumpen la parábola ideal curvada hacia arriba. La cesura temporal interviene, en el primer factor, entre 1941 y 1942 para ambos emisores. La cesura de 1941 que parece prematura con respecto a las mutaciones ideológicas, si aceptamos que el impacto del cambio ideológico debería dejarse notar, como lo sugieren numerosos estudios sobre diferentes *corpus*, con posterioridad a los acontecimientos y que no podría, por tanto, producirse antes de finales de 1942. Ello induce a pensar que un acontecimiento de la magnitud y duración de la guerra civil haya podido reflejarse en las prácticas discursivas hasta fechas muy posteriores a su conclusión extralingüística.

Hasta aquí hemos examinado el parentesco y antagonismos de los textos de *Ya y Arriba* sobre la base de sus perfiles léxicos. Antes de pasar a examinar el vocabulario que ha caracterizado los reagrupamientos señalados por el análisis, podemos puntualizar algunos de los interrogantes que se perfilan a partir de estos resultados.

Se plantea, en primer lugar, la pregunta clásica con respecto a la evolución, es decir, cuál es el vocabulario responsable de la misma. Qué formas caen en desuso y cuáles son aquellas que vienen a ocupar, cuantitativamente, su lugar. Qué formas caracterizan las etapas originales delimitadas por los resultados. Más concretamente se plantea saber por una parte, cuál es el vocabulario responsable de la cohesión de los dos diarios durante la primera etapa.

Trataremos de examinar si se trata, masivamente, de un vocabulario ligado a los temas comentados, lo cual no sería extraño dadas las normas de constitución del *corpus*, o se trata de indicios de funcionamientos discursivos ligados a procesos retóricos o a mecanismos enunciativos vinculados, por ejemplo, al género.

Por otra parte, en el plano cronológico, se trataría de ver cuáles son las formas más contributivas que singularizan los bloques temporales y a qué vocabulario corresponden las anomalías evolutivas de 1943, 1942 y 1940.

En espera de otro análisis que vengan a precisar y profundizar estos primeros resultados, estudiaremos los datos relativos al parentesco de las formas gráficas constitutivas del *corpus*.

---

<sup>4</sup> A. SALEM: «Approches du temps lexical. Statistique textuelle et séries chronologiques». *MOTS*, 17, octobre 1988, p. 105 a 144.

### 3.1.2. Las formas

Para el análisis de las formas, su proyección sobre el plano resulta excesivamente abigarrada (gráfico 2); la máquina proporciona el inventario de las formas más contributivas (los elementos línea que alcanzan un umbral mínimo en su contribución al factor) al análisis, acompañadas de índices numéricos. Para el primer factor, los resultados oponen un conjunto de formas características del principio a otro grupo de formas sobreempleadas al final del período. Reproducimos, a continuación (cuadro 2), un fragmento del mismo que comporta índices numéricos relativos al análisis (contribuciones absoluta y relativa), a la frecuencia de las formas y su repartición. El estudio de los resultados correspondientes a los otros factores desvelan nuevas oposiciones que nos permitirán aproximarnos a las irregularidades evolutivas que intervienen dentro de los dos bloques temporales mayores contribuidos por el análisis. Abordaremos sucesivamente las características de las formas más contributivas al análisis y los avatares de las dos etapas delimitadas por el análisis.

Las formas más contributivas responsables de la tipología anteriormente descrita son, masivamente sustantivos (54,42 %). Los adjetivos suman el 18,08 %, los verbos el 10,24 % y las llamadas formas funcionales, el «resto», 16,26 % entre las cuales los pronombres personales son en cualquiera de sus formas muy escasos. Ello supone, con respecto a los textos franceses sometidos al AFC, notables diferencias, ya que las tipologías se construyen a menudo a partir de la capacidad discriminante de las partículas lingüísticas ligadas a las personas del discurso, que son formas frecuentes y fluctuantes en los *corpus*. Se produce, por tanto, una especial orientación de las disimilitudes hacia categorías relacionadas con la enunciación cuya específica inscripción en la linearidad son producto de la lengua (en este caso concreto conformadas por el sistema pronominal y verbal). En español, dado el carácter sintético de las formas verbales, la superficie, segmentada, en formas mínimas que son sometidas posteriormente al análisis, refleja, en sus resultados, la presencia de materiales gramaticales diversos.

#### 3.1.2.1 El vocabulario del primer momento

Las proporciones que hemos mencionado anteriormente se hayan ligeramente alteradas a cada lado de los ejes. El principio del período (39-41) cuenta con una mayor proporción de sustantivos (58,43 por 100), en detrimento de los verbos que cuentan con el 12,04 por 100 en el segundo período.

La etapa más cohesionada, 1939-41, se caracteriza por la tendencia al superempleo de un conjunto de formas entre las que constatamos, en primer







311	italia	1.142	322	108	12	0	5	1	0	0	0	0	0	0	0	0	5	1
270	ejército	1.098	594	108	13	4	2	3	0	0	0	0	0	0	0	2	1	1
191	partido	1.090	550	148	18	4	5	0	0	0	0	0	0	0	0	4	3	0
423	glorioso	1.023	341	65	9	4	1	0	0	1	1	0	2	0	0	2	0	0
335	enemigo	994	475	75	11	2	2	1	1	0	0	0	0	0	0	0	4	1
424	grandeza	888	206	49	9	5	1	0	0	0	1	0	0	0	0	0	0	0
298	casi	879	364	64	12	0	3	0	1	0	0	0	0	0	2	3	3	3
90	victoria	817	549	166	36	11	0	4	0	1	2	0	0	3	5	5	5	5
290	sacrificio	806	376	58	13	3	0	0	0	2	0	0	0	3	2	1	1	1

Num: número de orden jerárquico de la forma.

Fl: valor del factor.

cos2: contribución relativa de la columna K<sub>hi2</sub>.

CPF: contribución relativa de la columna al factor.

FT: frecuencia total.

lugar, la presencia masiva de términos relacionados con el tema del combate, tales como

frente, guerra, lucha, combate, conquista, revolución, armas,  
ejército, paz, vencer...

que se eclipsan al final del período. Señalemos, asimismo, la fuerte presencia de términos ligados al aparato designativo. Destacan las referencias a Falange y sus organizaciones (juventudes, sindicatos, etc.), a España, Dios, ejército y las designaciones individuales y colectivas pertenecientes a los paradigmas de amigos / enemigos (*Alemania, Italia, alemán, enemigo /s, masoneria, roja, traidores...*)

Este primer período reúne también la presencia importante de lo que podemos llamar un vocabulario expresivo (en el sentido de la función expresiva de Jakobson), ligado al vocabulario de la lucha y a los sentimientos

angustiada, heroísmo, sacrificio/s, difícil/es, peligroso, dolorosa,  
delito/s, júbilo, amor, odio, alegría, santa, maldita, traidores,  
esperanza, anhelo, emoción, ingente, gran, amenaza, gratitud...

Las formas verbales reúnen verbos auxiliares (*es, fueron, ha, habido*), y escasos modales. Entre ellos encontramos alguna forma en primera persona de plural (*sabemos, estamos*) y, por otra parte, verbos de acción en tercera persona del singular e infinitivo (*pone, llega, manda, vencer, renacer, andar*). La exploración de los inventarios de formas proporcionados por el ordenador para cada factor nos permite aproximarnos al vocabulario responsable de las irregularidades evolutivas. Durante el primer período, las formas «descolgadas» son aquellas que se encuentran más alejadas de la norma (el origen de los ejes), aquellas que se encuentran en conjunción con *Arriba-1940* (superposición de términos y emisor que constataríamos en una representación simultánea que reuniera los gráficos 1 y 2) y son, por tanto, específicas de ese año y ese diario. Se trata de

camisa, militantes, minorías, fascista, europeo, Reich, andar, creído,  
comprenden, abiertos, angustias, tentación, diaria, estaban, Führer,  
Hitler, sueño, etc.



nados con la narración de los acontecimientos y la sucesión temática de los mismos (*realidad, situación, posición, relación, carácter, diversos, constituye, etc.*).

La categoría verbal concentra curiosamente numerosas formas pertenecientes a la declinación del verbo poder, ausentes en el primer período que acompañan a menudo verbos de habla (*podrán quedar desmentidas*).

En lo referente a las formas funcionales, advertimos, en la primera etapa, algunas partículas lingüísticas relacionadas con la primera persona del plural, lo cual, junto a las formas verbales de misma persona pueden constituir indicios de una más frecuente aparición del enunciador en el texto. No obstante, el carácter sintético de las formas verbales no autoriza, en español, a suponer directamente esta presencia. Señalemos, sin embargo, que un estudio posterior de los mecanismos enunciativos ha confirmado este fenómeno.

Advertimos, por otra parte, la presencia de varias formas de negación *ni* y *nadie*, a las que se une *no*, como forma menos contributiva pero también sobreempleada en el segundo período. El examen de los contextos aporta datos interesantes sobre el carácter de las negaciones aisladas por el análisis. Sin pretender por ello —dada la complejidad de formas que toma, en castellano— adelantar conclusiones sobre la negación en sus diferentes modalidades en el conjunto del *corpus*, si los retenemos como indicios de mecanismos discursivos con una considerable presencia en este segundo período. Observamos que *nadie* (cuya concordancia reproducimos en el cuadro 3), preponderantemente presente en esta segunda etapa, es a veces sujeto de un verbo de habla o de opinión.

«Nadie acusará de excesiva hojarasca retórica a esta definición de Arrese, en la cual se sintetiza [...] el problema de las provincias». (*Arriba*, 23 dic. 1943)

denotan la presencia, en el discurso, de otro discurso especialmente evocado para ser rebatido; son respuestas a un discurso del que esta prensa se defiende. Otras formas de negación acompañan términos políticos como libertades o democracia en sus definiciones.

«...el orden público no es una excepción en la vida de una colectividad, es...» (*Arriba*, 27 oct. 1945).

«...una mal entendida democracia que no era [...] más que un desbordamiento...» (*Ya*, 12 oct. 1945).

CUADRO 3  
CONCORDANCIA DE «NADIE» (Ya y Arriba, 1939-1945)

(Ordenación cronológica)

nadie

140060999	n llegado a esto? en realidad, no pueden culpar a nadie	sino a sí mismos. el período histórico que se cierra en agosto
140101203	¿cómo es tanto ma/s significativo cuanto que nadie	ignora la actual situación de nuestro visitante en el movimiento
140121322	¿trabaja o no? ¿cond/sima raíz/ cristiana para nadie	es desconocida. Y posiblemente nuestra tradición gramatical, tradici
241102165	dicionales de la mayoría de los teóricos de la ideología socialista.	ignore que esos son los términos para el conocimiento hoy obsoleto
141102233	- a atajarlo con la mayoría de los teóricos de la ideología socialista.	prefiere alinear exclusivamente y palabras precisas con que
142042378	mpuestos sobre el consumo de artículos, importa que nadie	será/ tan necio que suene con un tal estado pasivo que frente al
242042391	oral, por consiguiente, tan extraordinaria que nadie	ni aun el ma/s abastinado dentro de las naciones particulares,
142052428	o como norma. - pocas observaciones ma/s exactas que nadie	puede que pueda afirmar que "la guerra actual con sus características
142072459	que se contra el adoptamos posiciones antes que nadie	puede llamarse a ignominiosas que las reivindicaciones en
142113024	nductor gravemente consagrado de la Falange». nadie	es dado, por mucha libertad esperanzas derivadas de la consecuen
243033206	imperativo de los testamentos ni en la actitud de nadie	puede negarlo sin mala intención nuestro totalitarismo haya otor
243033214	supremo que ejerce el «Cuadrillo» de «espana». a nadie	debe pasar inadvertida la importancia de esta definición de
243123854	mejora ra/pidamente y de forma política y política que nadie	tan capacitado para hablar como nosotros, que cacion y trascen
144114870	enestar y al decoro. elecciones sindicales, para nadie	contradictoria que en la raíz y radical evolución se abraza
145044891	estructuraban la grandeza de «espana». pero a nadie	nos gana a nadie que resulta ma/s co/modo desconocernos y tergiver
145044893	desconocidos y tergiver los mismos argumentos que nadie	esta es la verdad y de ahí que pudiera ser, después de campana
245085071	o/n y nobleza, a privilegiar la verdad española. nadie	ni siquiera aquello que en el visto general se llamaba familia,
245085073	a todos, en tanto que no ha sido el mayor favor de nadie	discrepaba fundamentalmente del fondo del sistema y todos disc
245105186	s inquietudes, tanto económico como político. nadie	antes habí/a logrado reformar. so/lo queda, para los abastinado
145105221	o voluntad y la decisión a adoptar se interrumpe nadie	
145115340	rodigarlo por doquiera en aquellos tiempos en que nadie	
245115431	se puesto término a una situación de hecho que nadie	
27	OCURRENCES	

Indeterminados como *quienes* y determinantes como *algunos* acompañan categorías que el discurso trata de descalificar, adversarios políticos a los que no se designa de forma directa y a los que se refiere en la modalidad del «ya saben de quien se trata».

«Si algunos de nuestros compatriotas llegaran a imaginar que [...] pudiera alterarse lo que se haya [...] estatuido, se hallan en un completo error». (*Ya*, 10 nov. 42).

«Quienes no supieron establecer los más elementales fundamentos de convivencia». (*Arriba*, oct. 45).

La recurrencia de la conjunción *y* parece señalar una mayor presencia, en esta parte, de estructuras sintácticas de coordinación (que confirma la presencia de la forma negativa *ni*).

Los años 1942 y 1943 parecen deber su alejamiento de la norma en los factores 1 y 2 a un vocabulario que se opone en razón de la sucesión temática de los editoriales. Para 1942, el problema de las subsistencias, el desembarco de las tropas aliadas en el Norte de Africa, etc.:

abastos, comisaría, independientemente, reivindicaciones, imponen, totalitarios, reguladoras, alimentación, corredor, dará, democráticas, democrático, destrucciones, dispendios, intermediarios, irrenunciables, mercantiles, organizaciones, privadas, radica, representar, experiencia, salvarse, democrática, Marruecos, órgano, Africa.

Para Y43 y A43, las formas están vinculadas al tema de la educación, a la política familiar (*docente, universidad, universitarios, familiares, subsidios, mejoras*, etc.).

### 3.2. El escalograma 500 × 14

Una segunda experiencia llevada a cabo con las 500 formas más frecuentes (de frecuencia superior a 8) del *corpus* reproduce los primeros resultados. Constatamos, para el primer factor, las mismas asociaciones de



emisores y los mismos desequilibrios evolutivos. Con respecto al vocabulario, el cuadro 4 reproduce la lista de formas más contributivas son las siguientes:

#### CUADRO 4

Formas de  $F \geq 8$  más contributivas al AFC

##### **Vocabulario de principio de período**

homenaje, auxilio, Himmler, masonería, juventudes, Alemania, Italia, ejército, partido, glorioso, enemigo, grandeza, casi, victoria, sacrificio, difíciles, espíritu, línea, paz, sindical, les, acción, guerra, lucha, Falange, hoy, hasta, frente, nuevo, porque, ante, gran, nos, esta, ha, es, y, la.

##### **Vocabulario de final de período**

las, política, españoles, ni, vida, dentro, régimen, nadie, leyes, país, declaraciones, carácter, libertad, orden, través, quienes, pueda, aquellas, constitución, Cortes, cosa, pleno, excepción, bases, aquello, podrán, público, democracia, constitucional, administración, fuero, libertades, algunos, garantías, local, previsión.

#### CONCLUSIÓN

Desde un punto de vista ideológico el vocabulario de la primera época es, como esperábamos, más «azul» (como se dice de los primeros años de la dictadura por su coloración falangista). La exploración de las concordancias confirma la participación de las formas superempleadas al principio del período en el tema de la justificación de la guerra civil, la glorificación de la victoria, pero también la «lucha sin tregua» y los valores falangistas por la construcción de un orden nuevo, el combate de España contra sus enemigos. La segunda etapa parece evocar, por el contrario, valores a primera vista más bien relacionados con el discurso liberal. Se trata de términos relacionados con los temas de «libertad y orden» y de la legalidad institucional.

Tratemos, finalmente, de resumir los datos procedentes del análisis de la muestra de prensa en esta primera aproximación a su «historia cuantitativa» para acercarnos al funcionamiento discursivo de las secuencias textuales. Durante la primera etapa, caracterizada por la presencia preponderante de un vocabulario de lucha, en el que la expresividad ocupa un lugar importante, la cohesión constatada estaría relacionada con esta específica repetitividad. Dentro de este periodo, destaca la tendencia de *Arriba*, durante 1940, a evocar el quehacer político de la revolución nacional-sindicalista y de Falange como formación extradiscursiva, cuya designación (al igual que las de sus organizaciones de masas) es cuantitativamente importante únicamente en esta primera parte. A partir de 1942, la ausencia de esta repetitividad homogeneizadora repercutiría en una mayor sensibilidad de la caracterización al impacto temático. Tomaría pues su relevo la proliferación de un vocabulario más marcado por la sucesión de los acontecimientos. El marco general ha dejado de ser el del combate para ser el institucional, el político, el mundial. La coherencia final estaría debida a la proliferación de la estructura de coordinación, a la presencia de términos metadiscursivos y de la negación que parecen denotar, por una parte, una etapa de construcción referencial y, por otra, un discurso más polémico y argumentativo.

Podemos, a partir de los resultados, formular una hipótesis discursiva que trataremos de poner a prueba en el resto del *corpus*. A la primera época, caracterizada por un vocabulario que describe el combate pasado y presente contra el «enemigo de España», sucedería una segunda etapa caracterizada por otro tipo de relación con el adversario, centrada en una mayor beligerancia discursiva. El combate no sería esta vez evocado o «representado», sino que ocurriría en el propio tejido del texto y se traduciría entre otras cosas, en la lucha por la apropiación, término a término, de un vocabulario ligado al discurso liberal. Sería por lo tanto, en este sentido, un periodo más activo en la lucha contra el discurso del «otro».