

Catacterísticas estilométricas del sustantivo y del verbo

Observaciones de dos textos salteños de la actualidad

Hiroto Ueda

1. Introducción

Soy profesor japonés de la lengua española retirado de una universidad de Tokio. Durante mi carrera profesional, me he dedicado casi exclusivamente al estudio cuantitativo del español, que aprendí en la universidad durante dos años. Mi campo de especialidad es análisis estadístico de la lengua y he seguido dando clases de estadística y español durante cuarenta años. En esta ocasión brindada por mi estimado amigo poeta Eduardo Ceballos, voy a explicar un método de procesamiento lingüístico de textos, con una aplicación a las preciosas expresiones del habla salteña recogidas sucesivamente en 2004, 2005 y 2008, después de la primera visita que hice en 1996, invitado por el profesor Francisco Fernández a dar un curso de dialectología española.

Para mí es raro que visite la misma ciudad cuatro veces en la vida, fuera del caso habitual de Madrid, donde tengo proyectos de investigación con mis colegas españoles. Pensando que Salta es una de las ciudades más lejanas desde Tokio, casi antípoda, la frecuencia de mis visitas resulta todavía más raro. Lo raro de mis viajes se explica por lo mucho que me ha gustado por la amabilidad con la que me recibieron personas inolvidables en toda la vida. Conservo mucho cariño a todos ellos.

Ahora bien, aspectos sentimentales aparte, voy a estudiar lo más objetivamente posible dos textos salteños, unos de explicaciones de sitios visitados en las dos ocasiones y otros del libro obsequiado por Eduardo. En diciembre de 2004, pleno verano en Argentina, y seguidamente en septiembre de 2005, primavera, junto con los profesores y estudiantes de la Universidad Nacional de Salta, María Victoria Lladós, Francisco Fernández, Héctor Rodríguez, Cristóbal Fernández, recorrí distintos lugares de las provincias de Salta, Jujuy y Tucumán para conocer su cultura, naturaleza y la vida de la gente de cada localidad. Posteriormente edité los videoclips junto con sus transcripciones y algunos ejercicios. Están divididos en dos partes con 7 capítulos: PRIMERA PARTE: 1. De Jujuy a La Quiaca / 2. De Jujuy a Tartagal / 3. De Salta a Cafayate / 4. De Cafayate

a Salta; SEGUNDA PARTE: 5. Salta en fiesta / 6. Tucumán / 7. Volver a Salta. Los dos viajes fueron maravillosos. Se titula *Norte de Argentina*¹.



Chacarera en San Carlos, 2004

Con un intervalo de tres años, en septiembre de 2008, primavera en Argentina, visité de nuevo la ciudad de Salta, donde recibí la amable colaboración de la gente de la ciudad y del campo. En un tiempo libre, me permití hacer una entrevista a Eduardo, quien nos explicó sus inquietudes sobre el pasado y el presente de nuestro planeta. Nos hizo el favor de leer toda su obra: *Por amor a la vida*.



Eduardo Ceballos y su hija Viviana, 2008

Cito la introducción de Viviana Ceballos que presentó en el centro de la ciudad²:

POR AMOR A LA VIDA es un trabajo literario que tiende a generar conciencia para racionalizar la administración de los terrenales recursos. Para ello, el autor viaja por todo el espacio de la historia natural, desde el Big Bang hasta el presente de climas complicados. Un recorrido por la materia que viene desde lejos marcando distancias siderales e influencia en la tierra con su carga de neutrinos. / Luego, el asombrado mundo de la interioridad, donde se construye la personalidad y el yo, el núcleo de uno mismo. En ese camino sin sombras compara las células con los

¹ <https://spanish.ecc.u-tokyo.ac.jp/aries/argentina>

² <https://spanish.ecc.u-tokyo.ac.jp/aries/amorvida>

planetas y descubre similitudes que emocionan. / Luego de caminar por el espacio y por la carne, sale por los balcones de sus ojos a mirar el mundo circundante en el que comparten los hombres el paisaje. Todo está relacionado y se contiene. / La vida se reparte por el aire, por el agua, por la tierra. Los hombres construyen y destruyen, matan y viven, falta un poco de orden y equilibrio. / Es necesario devolver a las próximas generaciones este habitáculo terrenal sano. Las contaminaciones están enfermando la tierra y producen cambios que provocan muerte. / Se trabaja para generar conciencia POR AMOR A LA VIDA.

Los textos reunidos en los tres viajes fueron transcritos por los presentadores mismos, mientras que el texto del libro fue concedido por el autor. Al leerlos se observan diferencias estilísticas notables. Los primeros textos de *Norte de Argentina* son realizados con el objetivo de presentar la vida, cultura y naturaleza de la tierra salteña y sus alrededores con un estilo ameno y familiar, mientras que el libro del poeta, *Por amor a la vida*, instruye la ciencia y filosofía al lector con un tono elevado e inteligente³.

2. Lematización

A continuación voy a explicar brevemente el sistema de lematización, que acabo de construir en el sitio de la Universidad de Tokio. Según Gómez Díaz (2005: 117)⁴, la “lematización” es “la acción de extraer la esencia, es decir, el tema o lema de una palabra”. El término “lema (...) se trata de una etiqueta introducida por aplicaciones informáticas para hacer referencia a la base sobre la cual las distintas formas flexivas actúan”. En cuanto a la lematización para el español, la misma autora distingue entre la lematización flexiva y la derivativa (id: 152), de las cuales me he limitado a llevar a cabo solo la primera. De esta manera llegamos a las formas representativas que exponen los diccionarios en forma masculina singular de sustantivos, adjetivos, artículos, pronombres y posesivos, y en forma infinitiva de verbos. Por ejemplo, las formas *blanco*, *blancos*, *blanca*, *blancas* corresponden al lema «blanco»; y *como*, *comes*, ..., *comiendo*, *comido*, *comer*, al lema «comer».

Por otra parte, nos interesan también la categoría gramatical del lema y la información morfológica de la forma, por ejemplo, *blancos* = [adjetivo], {masculino. plural}, *como* = [verbo] {indicativo, presente, primera persona singular}. La misma forma

³ He utilizado *Norte de Argentina* en los cursos del primer y segundo cursos de la lengua española y *Por amor a la vida*, en el tercer y cuarto cursos de la misma.

⁴ Gómez Díaz, Raquel. 2005. *La lematización en español: una aplicación para la recuperación de información*. Gijón.

como puede representar también la conjunción o el relativo, de modo que hay que desambiguarla entre las tres categorías gramaticales, verbo y conjunción, de acuerdo con el contexto. Para ello, he pedido ayuda a mi colega español, al profesor Antonio Moreno Sandoval de la Universidad Autónoma de Madrid, quien generosamente me ha ofrecido gran cantidad de datos previamente lematizados de manera automática en su propio sistema Grampal y luego corregidos manualmente⁵. Son de 802.000 formas en secuencia con su lema, categoría gramatical e información morfológica correspondientes. He reeditado los datos para llegar a la forma conveniente para el propósito de investigación:

CATEGORÍA GRAMATICAL:

a	adjetivo
ab	abreviación
ar	artículo
av	adverbio
c	conjunción
cl	clítico
d	demonstrativo
e	extranjerismo
it	interjección
n	sustantivo
nm	numeral
nn	número
np	nombre propio
p	preposición
pn	pronombre
ps	posesivo
q	interrogativo
r	relativo
v	verbo
z	puntuación, paréntesis, etc.

VERBO :

1s	primera persona singular
----	--------------------------

⁵ Moreno, Antonio, Guillermo de la Madrid, Manuel Alcántara, Ana González, José M. Guirao and Raúl De la Torre (2005) "The Spanish Corpus" en Emanuela Cresti and Massimo Moneglia (eds) *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*. John Benjamins Publishing Company.

1p	primera persona plural
13s	primera / tercera persona singular
2s	segunda persona singular
2p	segunda persona plural
3s	tercera persona singular
3p	tercera persona plural
cond	condicional
fut	futuro
ger	gerundio
i	indicativo
idf	pretérito indefinido
imp	imperfecto
inf	infinitivo
ipr	imperativo
pp	participio pasado
pre	presente
s	subjuntivo

NO VERBO:

c	común
dat	dativo
f	femenino
m	masculino
n	neutro
p	plural
ref	reflexivo
s	singular
sp	singular / plural

Con estos datos y nuestro sistema *LYNEAL* («*Letras y Números en Análisis Lingüísticos*»)⁶ se consiguen los datos lematizados de la siguiente manera:

⁶ <https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/>
<http://shimoda.llf.uam.es/ueda/lyneal/>

Input text

Hola, buen día. Yo soy María Victoria; voy a acompañar al señor Ueda y al profesor Francisco Fernández en esta travesía. Espero que la disfruten ... que disfruten el norte de la Argentina; lo que es Salta y Jujuy. ¡Sean bienvenidos! Estamos ahora... Vamos a ver ahora lo que es la gente de Jujuy; a qué se dedica, qué hace, cómo se maneja en transporte. Estos son colectivos; aquí la gente se desplaza desde sus hogares hasta el lugar de trabajo, cuando no poseen vehículo. Son muy económicos y prácticos para viajar en la ciudad. Bueno... Atrás mío tengo la plaza principal de San Salvador de Jujuy. Como podrán ver, la plaza simboliza el aire puro, la naturaleza, la libertad. Los espacios verdes acompañan a lo que es la Casa de Gobierno... al que dirige toda esta provincia. Allí se encuentra el Gobernador de esta ciudad reunido con sus diputados, concejales, resolviendo los problemas sociales de esta sociedad que crece día a día. Bueno... Continuamos acá conociendo lo que es Jujuy; ya adentrándonos en la zona comercial de esta provincia. Puede verse la cantidad y variedad de locales y cómo la gente no sólo de Jujuy, sino de Salta y también de Bolivia vienen a esta zona a comprar por su bajo costo. Pueden verse desde ropas, hasta revistas, comidas: todo tipo de negocios. Inundan la zona céntrica de la provincia. Esta es una comida del norte. Se llama tamal; como verán, está envuelta en chala... [Esto es tamal... Ah no: humita... humita...] Esta es la chala del choclo y lo que tiene adentro, es harina de maíz con queso. Una comida muy del norte. Bueno, esta es la otra comida regional, típica del norte... Se llama tamal... También está envuelta en esta chala... Y la diferencia con la humita es que este trae carne... de res... con papas. Algunos le ponen huevo y esta es salada, a diferencia de la humita que puede ser dulce o salada. Esta es salada...
 Y de la hermosa ciudad en la que estábamos, pasamos

(Delete)

Output text

Form	Category	Lemma	Inflection
hola	it	hola	.
,	z	,	.
buen	a	bueno	m.s
día	n	día	m.s
.	z	.	.
yo	pn	yo	1s
soy	v	ser	i.pre.1s
María	np	María	.
Victoria	np	Victoria	.
;	z	;	.
voy	v	ir	i.pre.1s
a	p	a	.
acompañar	v	acompañar	inf
a	p	a	.
-l	ar	el	m.s
señor	n	señor	m.s
Ueda	np	Ueda	.
y	c	y	.
a	p	a	.
-l	ar	el	m.s
profesor	n	profesor	m.s
Francisco	np	Francisco	.
Fernández	np	Fernández	.
en	p	en	.
esta	d	este	f.s
travesía	n	travesía	f.s
.	z	.	.
espero	v	esperar	i.pre.1s
que	c	que	.
la	ar	el	f.s
disfruten	v	disfrutar	s.pre.3p
...	z
que	c	que	.
disfruten	v	disfrutar	s.pre.3p
el	ar	el	m.s
norte	n	norte	m.s
de	p	de	.
la	ar	el	f.s
Argentina	np	Argentina	.

(Delete)

[Run-1] Lemmatize text

[Run-2] Make lemma data

 Information Attribute

(Delete all)

Finish! 10 / 10 : 22293 (436 m.s.)

[\[Top page\]](#)<https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/etc/lemma.htm>

Nuestro método no adopta el modelo de procesamiento morfológico (Moreno Sandoval, 1998: 94-95, 2019: 145-170; Ueda 2016)⁷, sino el modelo probabilístico (Moreno Sandoval, 2019, 81-143), simplificado en forma de secuencias continuas de tres palabras. Al recibir el texto, el sistema busca la existencia de las tres palabras continuas en los

⁷ Moreno Sandoval, 1998, *Lingüística computacional*. / Moreno Sandoval, 2019, *Lengua y computación*. Madrid. Síntesis. / Ueda, Hiroto. 2018. "Tratamiento lingüístico y matemático de textos digitales españoles. Presentación del Programa LEXIS-web", Actas del IX Congreso de la Asociación Asiática de Hispanistas (Bangkok, 2016), pp. 617-630. http://www.sinoele.org/images/Revista/17/monograficos/AAH_2016/AAH_2016_hiroto_ueda.pdf

datos previamente almacenados. Cuando se encuentran dos o más casos, se selecciona el caso dotado de más probabilidad. Al no encontrar la secuencia de tres palabras, el sistema procede a la búsqueda de dos palabras continuas. Al no encontrar la secuencia de dos palabras continuas, procede a la búsqueda de una palabra. Se aprovechan los datos de lema, categoría e información que se hallan en la memoria directa inmensa de las posibles formas correspondientes. Pocas veces ocurre que no se encuentra la palabra en los datos. En ese momento, el sistema devuelve el signo de asterisco (*), lo que se aprovecha con las anotaciones manuales para los procesamientos siguientes. De esta manera el sistema se enriquece en el recurso de palabras en forma de datos morfológicos, con la frecuencia y probabilidad⁸.

3. Sustantivo y verbo

Según los estudios anteriores, el uso relativo de sustantivos respecto a los verbos constituye un rasgo estilístico importante en distintas lenguas tanto europeas como japonesa. Moreno Sandoval y Guirao (2006: 214) explican que el sustantivo y adjetivo aparecen con mucha más frecuencia en los textos escritos que en los hablados⁹. La misma tendencia se observa en la lengua japonesa (National Institute for Japanese Language and Linguistics, 1955, 140-141)¹⁰. Veamos la diferencia entre el estilo coloquial familiar (*Norte de Argentina*) y el estilo escrito elevado (*Por amor a la vida*):

Norte de Argentina

Estamos ahora... Vamos a ver ahora lo que es la gente de Jujuy; a qué se dedica, qué hace, cómo se maneja en transporte. Estos son colectivos; aquí la gente se desplaza desde sus hogares hasta el lugar de trabajo, cuando no poseen vehículo. Son muy económicos y prácticos para viajar en la ciudad.

Bueno... Atrás mío tengo la plaza principal de San Salvador de Jujuy. Como podrán ver, la plaza simboliza el aire puro, la naturaleza, la libertad. Los espacios verdes acompañan a lo que es la Casa de Gobierno... al que dirige toda esta provincia. Allí se encuentra el Gobernador de esta ciudad reunido con sus diputados, concejales,

⁸ cf. McEnery, Tony & Wilson Andrew. 1996. *Corpus Linguistics*. Edimburg University Press. pp. 119-120.

⁹ Moreno Sandoval, Antonio y Guirao, José M. 2006. "Morpho-syntactic Tagging of the Spanish C-ORAL-ROM Corpus. Methodology, Tools and Evaluation", in *Linguistic Informatics and Spoken Language Corpora. Contributions of Linguistics, Applied Linguistics, Computer Sciences*, Tokyo University of Foreign Studies, pp. 199-218.

¹⁰ National Institute for Japanese Language and Linguistics, 1955, *Danwago no zittai* (Realidad de las palabras en discurso).

resolviendo los problemas sociales de esta sociedad que crece día a día.

Por amor a la vida

+ ¡Hola! ¿Cómo estás? ¿Cómo ya no me conoces?

- Pero, ¿quién eres?

+ Soy tu otro yo, el alter ego. Hace más de sesenta años que compartimos la existencia.

- Entonces, ¿somos la misma cosa?

+ Efectivamente somos la misma persona... Trabajemos juntos en un monólogo, o un diálogo conmigo mismo, o un soliloquio, que lo quería ofrecer como un testimonio de mi paso por la vida...

- ¿Un mono qué?

+ Un monólogo que devuelva parte de lo mucho recibido.

- ¿Tiene que ver con el monotributo?

+ No. El monólogo es un caudal de recuerdos, emociones que vienen de la memoria, para mostrar el paisaje interior y esas personales fotos que uno saca con el alma de ese mundo circundante.

- ¿Y yo de qué juego? ¿De mono solo?...

+ No, porque eres parte de mi mismo, así que juntos abriremos el cofre donde están grabados los distintos momentos de la existencia.

- Tu pedido es una orden. Comienza como quieras...

+ Tomo las llaves, abro el íntimo rincón de mi yo, y desde las ventanas de mis ojos observo el mundo. Percibo su enfermedad y eso mueve mi sensibilidad y surge la necesidad de expresarme por la salud de la tierra, POR AMOR A LA VIDA.

El sistema devuelve las siguientes tablas:

N	Forma	Categoría	Lema	Inflexión
41	lo	ar	el	n
42	que	r	que	.
43	es	v	ser	i.pre.3s
44	Salta	np	Salta	.
45	y	c	y	.
46	Jujuy	np	Jujuy	.
47	.	z	.	.
48	i	z	i	.

49	sean	v	ser	s.pre.3p
50	bienvenidos	a	bienvenido	m.p
51	!	z	!	.
52	estamos	v	estar	i.pre.1p
53	ahora	av	ahora	.
54	...	z
55	vamos	v	ir	i.pre.1p
56	a	p	a	.
57	ver	v	ver	inf
58	ahora	av	ahora	.
59	lo	ar	el	n
60	que	r	que	.
61	es	v	ser	i.pre.3s
62	la	ar	el	f.s
63	gente	n	gente	f.s
64	de	p	de	.
65	Jujuy	np	Jujuy	.

NORTE DE ARGENTINA

N	Forma	Categoría	Lema	Inflexión
14	¡	z	¡	.
15	hola	it	hola	.
16	!	z	!	.
17	¿	z	¿	.
18	cómo	q	cómo	.
19	estás	v	estar	i.pre.2s
20	?	z	?	.
21	¿	z	¿	.
22	cómo	q	cómo	.
23	ya	av	ya	.
24	no	av	no	.
25	me	cl	me	1s
26	conoces	v	conocer	i.pre.2s
27	?	z	?	.
28	-	z	-	.
29	pero	c	pero	.

30	,	z	,	.
31	¿	z	¿	.
32	quién	q	quién	s
33	eres	v	ser	i.pre.2s
34	?	z	?	.
35	+	z	+	.
36	soy	v	ser	i.pre.1s
37	tu	ps	tu	c.s
38	otro	a	otro	m.s
39	yo	pn	yo	1s
40	,	z	,	.
41	el	ar	el	m.s
42	alter	e	alter	.
43	ego	e	ego	.
44	.	z	.	.

POR AMOR A LA VIDA

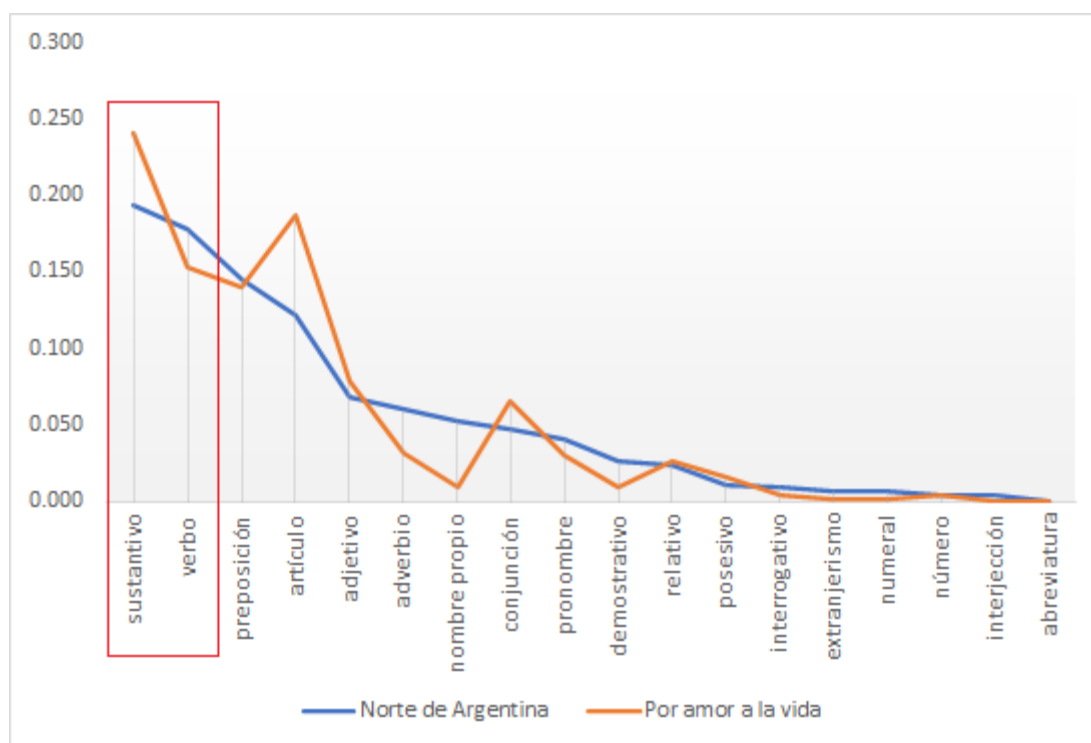
A partir de estas tablas, que contienen 22 118 palabras en *Norte de Argentina* y 4 591 palabras en *Por amor a la vida*, he calculado las frecuencias absolutas y relativas de todas las categorías gramaticales que hay en ambos textos:

s

<i>Categoría</i>	Frecuencia absoluta		Frecuencia relativa	
	<i>Norte de Argentina</i>	<i>Por amor a la vida</i>	<i>Norte de Argentina</i>	<i>Por amor a la vida</i>
n	940	3133	0.241	0.194
ar	727	1971	0.187	0.122
v	596	2868	0.153	0.178
p	546	2348	0.140	0.145
a	305	1103	0.078	0.068
c	255	759	0.065	0.047
av	125	970	0.032	0.060
pn	119	661	0.031	0.041
r	104	381	0.027	0.024
ps	62	168	0.016	0.010
d	37	416	0.009	0.026
np	36	850	0.009	0.053
q	15	155	0.004	0.010

nn	14	72	0.004	0.004
e	7	113	0.002	0.007
nm	7	108	0.002	0.007
it	2	66	0.001	0.004
ab		7	0.000	0.000
intj		2	0.000	0.000
Total	3897	16151	1.000	1.000

El siguiente gráfico es de las frecuencias relativas de las categorías más importantes:



Llaman la atención las primeras dos categorías gramaticales, el sustantivo y el verbo, por su alta frecuencia en general y el cambio de orden entre *Norte de Argentina* (estilo coloquial familiar) y *Por amor a la vida* (estilo escrito elevado). Es decir, el sustantivo se presenta más en el estilo escrito elevado, mientras que el verbo se utiliza más en el estilo hablado coloquial. La diferencia de la frecuencia de sustantivo entre los dos textos dentro de la totalidad es estadísticamente significativa ($\chi^2 = 43.258$, $p < .01$). La diferencia de la frecuencia del verbo, lo es también ($\chi^2 = 13.331$, $p < .01$). ¿A qué se debe esta diferencia?

4. Reflexiones

En múltiples estudios anteriores está demostrada la relación antitética que hay entre la frecuencia del sustantivo y la del verbo. Cuando uno aumenta en el texto, el otro tiende a disminuir en el mismo texto, lo que no ocurre por ejemplo entre el sustantivo y el artículo, que correlacionan naturalmente por razones sintácticas: artículo + sustantivo. La frecuencia del sustantivo correlaciona con la del verbo de manera inversa.

Ouno (1956) hizo el recuento de las dos categorías en distintos géneros de la literatura clásica japonesa para comprobar la tendencia inversa: cuánto más el sustantivo, menos el verbo, y viceversa¹¹. Hatano (1958: 142-143) calculó los usos de sustantivos y verbos en las obras literarias modernas de dos grandes novelistas japoneses, de los cuales uno utiliza más verbos y produce el efecto dinámico en oraciones largas, mientras que el otro está contrariamente caracterizado por el uso frecuente del sustantivo, cuyo efecto es estático en oraciones cortas¹².

El sustantivo y el verbo son diferentes en su distribución de frecuencia. Tanto es así que en los análisis multivariados suelen situarse en los dos puntos opuestos con respecto al centro de la escala. Esto se puede comprobar en el análisis factorial de Biber (1988: 89) sobre dos estilos, hablado y escrito del inglés; y en el análisis de correspondencia de transcripciones de diferentes hablantes ingleses (Brezina, 2018: 201)¹³.

Por mi parte, he aprovechado los datos ofrecidos por el profesor Antonio Moreno Sandoval, mencionados en la sección 1. Se tratan de cinco conjuntos de textos:

C-ORAL: Habla espontánea, informal.

MAVIR: Charlas profesionales, conferencias universitarias de temas de informática.

OCU: Texto médico, escrito por periodista para el público general.

MdC: «*Museo de Cera*», poemario de poeta español, José Álvarez.

Otro médico: Página web de consulta, escrita por médicos.

¹¹ Ouno, Susumu. 1956. “Kihongoi ni kansuru nisanno kenkyuu (Estudios sobre el vocabulario fundamental)”, en *Kokugogaku*, vol. 24.

¹² Hatano, Kanzi. 1958. *Kotoba to bunshoo no shinrigaku. (Psicología de palabras y oraciones)*. Tokio. Sinchoosha.

¹³ Biber, Douglas. 1988. *Variation across Speech and Writing*. Cambridge University Press. / Brezina, Vaclav. 2018. *Statistics in Corpus Linguistics. A Practical Guide*. Cambridge University Press.

Frecuencia absoluta	C-ORAL	MAVIR	OCU	MdC	Otro médico
Sustantivo	46 232	13 564	76 165	833	6 501
Verbo	66 116	14 591	50 759	539	3 349
Total	112 348	28 155	126 924	1 372	9 850

Frecuencia relativa	C-ORAL	MAVIR	OCU	MdC	Otro médico
Sustantivo	.412	.482	.600	.607	.660
Verbo	.588	.518	.400	.393	.340
ISV	-.177	-.036	.200	.214	.320

La última fila de la segunda tabla, ISV representa el Índice de Sustantivo y Verbo, que se calcula en la siguiente ecuación:

$$ISV = \frac{S - V}{S + V}$$

donde S es la frecuencia del sustantivo y V es la del verbo. Cuanto más es S, el ISV se aproxima al máximo 1 y cuanto más es V, el ISV se aproxima a -1. El ISV se vuelve 1 cuando V = 0, se vuelve -1 cuando S = 0; y presenta el valor 0 cuando S = V. Las tablas anteriores confirman la tendencia general del uso del sustantivo con respecto al del verbo, paralelo a la formalidad del estilo. A mi modo de ver, la escala de ISV representa la escala que recorre de lo coloquial a lo formal en el recorrido del mínimo -.177 (C-ORAL) al máximo .320 (Otro médico).

En la clase de la lengua española dirigida a los universitarios japoneses, a veces me encuentro con las traducciones japonesas del texto español difíciles de entender. Entonces recomiendo a los alumnos que utilicen verbos japoneses para los sustantivos españoles cuando la traducción directa resulta incomprensible. De esta manera, el estilo con más verbos se presenta más ameno, ampliado, explicativo, fácil de entender. Al contrario, cuando escribimos artículos científicos, preferimos buscar los sustantivos precisos para expresar el asunto exacto en cuestión. El estilo con más sustantivos se pone conciso. La diferencia parece deberse a la impresión de que el verbo conjugado se desarrolla en forma de cláusulas (*cuando viajamos a Argentina, ...*), mientras que el sustantivo se condensa en forma de frase corta (*en nuestro viaje a Argentina*), dentro de una oración compuesta de cláusulas y frases¹⁴. De ahí que se presenten por un lado la

¹⁴ Cf. Miles, J. 1964. *Eras and Modes in English Poetry*. Berkeley.

imagen dinámico en movimiento del viaje parecida a la película en la cláusula verbal y, por otro, el concepto concreto estático a modo de la pintura en la frase sustantiva. La película lleva más información visual que la pintura, mientras que la pintura permite más libertad en la interpretación.

La razón de la correlación entre el estilo y la tasa de sustantivo en la lengua japonesa fue explicada por la dependencia del contexto por Kabashima (1954: 18)¹⁵. Según su investigación sobre distintos textos japoneses, el sustantivo aumenta a lo largo de la escala desde las conversaciones diarias, hasta los artículos del periódico, pasando por los diálogos en novelas, libros de filosofía y narraciones en novelas. Este orden obedece al grado de la dependencia del contexto de la situación. La comunicación en la conversación diaria, sin muchos sustantivos, depende de las expresiones faciales, gestos, situación, objetos y comprensión previa que hay entre los hablantes, de los que carecen los artículos periodísticos, donde no pueden elidir los sustantivos que conllevan informaciones necesarias.

Yo añadiría otro factor importante: simpatía o confianza entre los hablantes. Sin ella, la comunicación resultaría sumamente difícil. Quizá por esta razón, en el estilo coloquial utilizamos más pronombres y demostrativos en lugar de sustantivos. En nuestra conversación diaria con la simpatía o confianza, podemos prescindir de los sustantivos, cosa que no ocurre en las descripciones periódicas, que no presuponen el compartimiento de los elementos arriba mencionados, ni la simpatía interpersonal desde el principio.

En la mayoría de estudios estilométricos, se analizan palabras frecuentes, especialmente en la atribución de la autoría de textos (Cerezo Soler, Juan y Calvo Tello, José. 2019. Pp. 235-236)¹⁶. En los estudios estilísticos de las obras literarias, se incluyen las categorías gramaticales dentro de multitud de otros rasgos cuantitativos (Yasumoto 1960, Kabashima y Sugaku 1965)¹⁷. Pienso que las unidades tan fundamentales como las categorías gramaticales son útiles para estudiar cuantitativamente las lenguas por sus características estilísticas comunes entre ellas. Es realmente sorprendente que las unidades lingüísticas no sean simplemente funcionales en el plano gramatical, sino

¹⁵ Kabashima, Tadao. 1954. "Gendaibun ni okeru hinshi no hiritu to sono zougen no youin ni tsuite" (Sobre la ratio de categorías gramaticales en el japonés actual y las causas de su aumento y disminución). *Kokugogaku* vol. 18, pp. 15-20.

¹⁶ Cerezo Soler, Juan y Calvo Tello, José. 2019. "Autoría y estilo. Una atribución cervantina desde las humanidades digitales. El caso de *La conquista de Jerusalén*", *Anales Cervantinos*, vol. LI, pp. 231-250.

¹⁷ Yasumoto, Biten. 1960. *Bunshou shinrigaku no shin ryouiki*. (Nuevos campos de la psicología de oración). Tokio. Sougensha. / Kabashima, Tadao y Sugaku, Akiko. 1965. *Buntai no kagaku (Ciencia estilística)*. Tokio. Sougousha.

también construyan directamente los rasgos estilísticos con la distinción de la familiaridad y la formalidad.

Por otra parte, ya para terminar el argumento, creo que en la actualidad, para llevar a cabo los estudios estilísticos, son necesarios tanto la reflexión cualitativa como el análisis cuantitativo, lo que era difícil en la época de Hough (1968), quien rechazó los métodos estadísticos de manera contundente¹⁸. Actualmente, sin embargo, estamos en condiciones de ampliar nuestro horizonte de ciencias humanas, ahora con los inmensos datos almacenados en forma de corpus lingüísticos y las nuevas herramientas de lematizadores y otros analizadores lingüísticos. Contando con ellos, deseo ver la verdad de la lengua, objeto de mi pasión de toda la vida.

¹⁸ Hough, Graham G. 1968. *Style and Stylistics*. Routledge.