

PHONETIC SIMILARITY BETWEEN JAPANESE AND SPANISH SYLLABIC SOUNDS

by Dr. Miguelina Guirao

ABSTRACT

This work is part of a study in progress intended to obtain evidence about phonetic similarities and differences between sounds of Japanese and Spanish speech. First a linguistic coincidence between both vocalic and consonantal systems is indicated. Then, the spectrographic analysis of Japanese vowels and subsequent comparison to the Spanish ones is presented. Segmental units—phonemes and syllables—were separated in acoustical groups and their perceptual parameters were examined. In a series of identification tests Spanish speakers participants were asked to listen to Japanese sounds and to transcribe them in their own alphabetic symbols. The main identification test consisted of one hundred basic syllabic sounds of the V, VV, CV and CVV type. Results showed that 81 of the syllables were well recognized. Percent of identification varied from 70 to 100%. The other 19 syllables rated from 30 to 60%. A correlation was also observed between acoustical characteristics of syllabic Japanese sounds and perceptual performance. Periodic sounds with harmonic components were better identified than aperiodic sounds. Syllabic segments of the type V, VV and CV gave higher percentage of identification than the type CVV.

Perceptual tests, are also in progress, in which Japanese-speaking listeners are presented with Spanish syllabic sounds. The sample includes V, VV, CV and CVC syllables and CVCV words. Japanese listeners were instructed to listen, to repeat for recording and to transcribe the sounds using their own symbols and also romanized letters. Preliminary results are showing that they are able to identify the sounds of the other language with equal good performance as the Spanish-speaking participants. Syllabic types V, CV and CVCV words reached almost 100%. The percentage were somewhat lower but still significant, for the other types of syllables. A more extensive study is envisaged to compare the more relevant prosodic features for both speech systems.

SEMEJANZA FONETICA ENTRE SONIDOS SILABICOS DEL JAPONES Y DEL ESPAÑOL

Dra. Miguelina Guirao

El tema de la comparación fonética entre los sonidos de la lengua japonesa y de la lengua española forma parte de uno de los proyectos de investigación que estamos desarrollando en el Laboratorio de Investigaciones Sensoriales del Consejo Nacional de Investigaciones Científicas y Técnicas. Nuestro interés se dirige principalmente a las estructuras básicas del habla: fonemas, sílabas y palabras. Estas unidades o segmentos de habla se analizan de acuerdo a parámetros fonéticos, fisiológicos, acústicos, estadísticos y perceptuales. En nuestro proyecto hacemos especial hincapié en la sílaba como unidad acústica y perceptual de la cadena hablada¹.

En esta exposición intentamos mostrar similitudes y diferencias en las estructuras básicas de cada lengua, independientemente de las reglas ortográficas y gramaticales. Por esa razón el estudio comparativo gira en torno de las características fonéticas y perceptuales de los sonidos que forman la secuencia de habla.

En líneas generales nuestro trabajo se distribuyó teniendo en cuenta tres aspectos fonológicos: En primer lugar tomamos nota de las coincidencias entre los fonemas que componen los respectivos sistemas vocálicos y consonánticos.

Luego hemos tratado de describir con técnicas experimentales las características físicas que tienen en común los sonidos y ruidos en la cadena acústica.

¹ M. Guirao, The syllable as an auditory signal. 10th International Congress on Acoustics. Sidney, Australia, 1980.

Por último hemos realizado pruebas perceptuales que nos permitieran evaluar el grado de parentesco fonético, que según suponíamos había entre las dos lenguas.

Similitud en los sistemas vocálicos y consonánticos

El japonés y el español comparten un sistema vocálico formado por cinco vocales bien diferenciadas /i, e, a, o, u/. Estos sonidos se pueden pronunciar en forma aislada y con diferentes funciones lexicales. En japonés la vocal *e* significa "mango" y "alimento", la *o* "cordón" y "cola", la *u* un "ave marina" y la *a* "mudo" y "casi". En español las vocales también tienen funciones gramaticales como la preposición *a* y las conjunciones *i e o y u*.

También observamos que con excepción de *e*, el orden de frecuencia de uso de las vocales es el mismo. La vocal *a* se usa más que *o* y esta más que *i*, siendo *u* la menos frecuente. En la tabla 1 insertamos los correspondientes porcentajes^{2 3}.

Tabla 1

Frecuencia de Uso de las Vocales

	Japonés	Español
/a/	13.3	12.4
/o/	10.8	9.8
/i/	7.3	7.2
/u/	2.9	3.8
/e/	4.6	14.5

² M. Guirao y A.M.B. Manrique, Fonemas, sílabas y palabras del español de Buenos Aires. *Filología*, 16:135-165, 1972.

³ M. Guirao y otros. *Estudio estadístico del español* (libro en preparación) 1984.

Veamos ahora el sistema consonántico. De nuestros 17 sonidos /p t k b d g m ñ r rr l s f x ʒ tʃ/ quince se pronuncian también en japonés. Las excepciones son /rr/ y /r/. Esta última suele considerarse como un fonema intermedio entre /r/ y /l/. Por su parte hay tres sonidos /ts/, /z/ y /ʃ/ que no existen en español aunque el último suele aparecer como una variación de /ʒ/ y de /tʃ/, en caracteres gráficos y *griega* y *ch* respectivamente.

Hemos notado además algunas semejanzas en relación con la sílaba. En nuestra lengua la sílaba constituye una forma básica formada por un núcleo vocálico al que le sigue o precede otra vocal o una consonante. El japonés también está compuesto de unidades regulares que se denominan "moras". Salvo algunas diferencias en el modo de segmentar las dobles consonantes y las terminaciones en *n*, la mora es prácticamente equivalente a la sílaba.

La combinación consonante-vocal (CV) constituye la sílaba típica japonesa. Este tipo silábico también predomina en español. En el cuadro Nro. 1 presentamos las 60 sílabas CV que son comunes a las dos lenguas.

En nuestro recuento estadístico el tipo silábico CV tuvo una incidencia del 58,6% el que le sigue, consonante-vocal-consonante (CVC) llega al 17,8%. Ambos sistemas tienen otras tres clases de sílabas que se usan con poca frecuencia: las que están formadas por una sola vocal (V), por dos vocales (VV) —la primera V suele llamarse "semivocal"— y por las que combinan una consonante con dos vocales (CVV). A estas clases el español agrega otras como CVVC, CCV y VC que sumadas no alcanzan al 10%³.

Con respecto a los fonemas que comienzan y terminan las palabras se dan también algunas coincidencias³. En ambos casos las consonantes que tienen mayor incidencia al comienzo de la palabra son /k/ y /s/. Le siguen aunque no exactamente en el mismo orden /m/ /n/ /t/ y /d/. En cuanto a los sonidos finales, para algunos lingüistas todas las palabras japonesas terminan en vocal⁴. Aun las terminaciones en *n* como en *hon* (libro), se consideran "vocálicas nasales". En español también hay una tendencia a terminar las palabras en vocal o en sonidos que, por su periodicidad, se emparentan

⁴ K. Naito, *Die Japanischen Phoneme. Proceedings of the 5th International Congress of Phonetic Sciences*. Munich, Alemania, 1964. S. Karger, Basilea, 1965.

	a ア	e エ	i イ	o オ	u ウ
p	pa パ	pe ペ	pi ピ	po ポ	pu プ
t	ta タ	te テ	ti チ	to ト	tu ツ
k	ka カ	ke ケ	ki キ	ko コ	ku ク
b	ba バ	be ベ	bi ビ	bo ボ	bu ブ
d	da ダ	de デ	di ヂ	do ド	du ヅ
g	ga ガ	ge ゲ	gi ギ	go ゴ	gu グ
m	ma マ	me メ	mi ミ	mo モ	mu ム
n	na ナ	ne ネ	ni ニ	no ノ	nu ヌ

Sílabas consonante-vocal del japonés y del español

/ɲ/	ña ニャ	ñe	ñi	ño ニョ	ñu ニユ
/l/	la	le	li	lo	lu
/r/	ra ㇰ	re ㇱ	ri ㇲ	ro ㇳ	ru ㇴ
/rr/	rra	rre	rri	rro	rru
/s/	sa ㇰ	se ㇱ	si	so ㇳ	su ㇴ
/f/	fa	fe	fi	fo	fu ㇰ
/ʒ/	ya ニャ	ye	yi	yo ニョ	yu ニユ
/x/	ja ハ	je ヘ	ji ヒ	jo ホ	ju
/tʃ/	cha ㇰャ	che	chi ㇰ	cho ㇰョ	chu ㇰユ

con las vocales. Según datos de nuestro trabajo estadístico, tomando muestra de hasta el 70% de las palabras que son diferentes y que aparecen con mayor frecuencia, los sonidos finales se distribuyen de este modo: 61,4% son vocales, 21,1% son las consonantes /n/ /r/ y /l/ y el resto, 16,3% corresponde a /s/, que no es un sonido periódico³.

Semejanza en las características acústicas de los sonidos

El primer trabajo de comparación que hicimos consistió en cotejar las características acústicas de los sonidos vocálicos. Como ya disponíamos de los registros espectrográficos de nuestras vocales⁵, procedimos a encarar el mismo estudio con las vocales japonesas. Las grabaciones se hicieron con la colaboración de once hablantes de ambos sexos estudiantes graduados y profesores del Departamento de Psicología de la Universidad de Hokkaido, Sapporo.

Una vez realizado el tratamiento acústico mediante el cual analizamos principalmente los componentes de frecuencia y amplitud en función de la duración, procedimos a efectuar la confrontación entre uno y otro grupo de registros. El cotejo reveló que el perfil físico de /i/ /e/ /a/ era esencialmente el mismo y el de /o/ muy similar aunque la /o/ japonesa tiene ciertas características de nuestra /u/. En cambio la /u/ japonesa no se equipara totalmente con la nuestra ya que comparte el formante de menor frecuencia (F_1) también con /i/ y /e/, y el formante que le sigue (F_2) con /o/⁷.

En vista de las similitudes que habíamos encontrado hasta el momento, procedimos a ordenar los sonidos del japonés siguiendo una clasificación acústica que ya habíamos propuesto para los sonidos del español de Buenos Aires⁶. Los dividimos en cuatro

⁵ M. Guirao y A.M.B. de Manrique, Identification of Argentine Spanish vowels. *Journal of Psycholinguistic Research* 4 (1): 17-25, 1975.

⁶ M. Guirao, Psychoacoustic correlates of phonetic features in Spanish sounds. *Speech Communication Papers*, 97th Meeting of the Acoustical Society of America, New York, USA:427-430.

⁷ M. Guirao, Similarity between Japanese and Spanish sounds. *The Study of Sounds*, 18 (Proceedings of the 3rd World Congress of Phoneticians, Tokyo, Japón, 1976): 211-216, 1978.

grupos: periódicos vocálicos, periódicos no vocálicos, pulsos y bandas de ruido. El primero reúne las cinco vocales que exhiben una clara composición armónica y con formantes bien definidos. Las consonantes del segundo grupo periódico /l/ /m/ /n/ /r/ presentan formantes de frecuencias relativamente bajas y fijas, con menor concentración de energía y menos diferenciados que los vocálicos. El tercer grupo lo forman seis pulsos que tienen un espectro irregular y una duración relativamente menor que los otros sonidos. Tres /p, t, k/ se caracterizan por presentar un cambio abrupto en la amplitud y los otros tres /b, d, g/ con cambios de amplitud más atenuados y con alguna periodicidad. El cuarto grupo /s, f, x, tʃ, ʒ/ prácticamente sin armónicos está compuesto por bandas de ruidos de diferentes frecuencias. En esta clasificación faltan las consonantes *rr* y *ñ* porque no encuadran totalmente en ninguno de los grupos descriptos. El sonido /rr/ es casi periódico y tiene picos en un extenso rango de frecuencias. El sonido *ñ* en símbolo fonético /r/ tiene características de sonido de fricción y de sonido periódico del tipo /n/ con la vocal /i/.

Por supuesto cuando oímos esos sonidos, ya sea aislados o en la secuencia de habla, estos no nos llegan como "formas acústicas" sino como "formas perceptuales" que en última instancia son las que fundamentan las distinciones fonéticas. Por esa razón hemos tomado especial cuenta de la relación que puede encontrarse entre una señal acústica y su correspondiente correlato psicoacústico o categoría perceptual. En general los sonidos que pertenecen a un mismo grupo acústico tienen también las mismas cualidades o categorías perceptuales. Así las cinco vocales se agrupan según la "cualidad tonal" y de acuerdo a los valores de amplitud y de frecuencia de los formantes se perciben como con mayor o menor sonoridad más o menos agudas o graves. Los sonidos periódicos no vocálicos tienen cualidad tonal más difusa y son menos compactos. Los pulsos de poco volumen y corta duración no alcanzan a definir una cualidad tonal propia sino generalmente la toman de la vocal. En contraste con el resto y especialmente con los vocálicos, tenemos los aperiódicos o ruidos cuya dimensión específica es la "aspereza" o "disonancia" como opuesto a consonancia⁸.

⁸ M. Guirao y C.R. Luis, Identification of Japanese syllables by Spanish-speaking listeners. *Journal of the Acoustical Society of Japan (E)*, 3 (1): 21-26, 1982.

Con respecto a estas dimensiones perceptuales debemos señalar que en ambas lenguas los sonidos armónicos tienen un notable predominio sobre los ásperos o disonantes. Veamos los datos estadísticos para el español. Entre los fonemas que se repiten con mayor frecuencia el 68% tienen cualidad tonal, el 48% de éstos son periódicos vocálicos y el 20% son periódicos consonánticos. Del resto el 19% son pulsos, tres de los cuales presentan alguna periodicidad y tres la adquieren, en cierto modo, porque se pronuncian junto con la vocal. Los cinco ruidos con el 11%, tienen una incidencia mucho menor y véase que de entre éstos el 9,5% corresponde a /s/ que es el menos áspero y de los otros cuatro, ninguno llega al 1%³. Aunque no se dispone de un recuento equivalente para fonemas japoneses podemos deducir de otros datos estadísticos⁹ que un alto porcentaje corresponde a sonidos periódicos. Por otra parte la gran riqueza de vocales en la sílaba dan al habla japonesa una cadencia armónica.

Pruebas perceptuales

Para verificar fehacientemente hasta qué punto los hablantes de español podían identificar sílabas japonesas, diseñamos varias pruebas perceptuales. En las primeras utilizamos las vocales japonesas con las cuales habíamos hecho el estudio espectrográfico⁷. Encontramos que /i/ /e/ /a/ se identificaron fácilmente y sin ningún error. La vocal /o/ también fue bien identificada pero los oyentes manifestaron que sonaba parecida a nuestra /u/. La /u/ en cambio fue reconocida pero más bien como a la /u/ francesa y como una "vocal indefinida"⁷.

Los pre-experimentos que habíamos realizado con sílabas CV nos animaron a continuar el estudio. Para la segunda serie de pruebas de identificación tomamos las 100 sílabas que generalmente se usan para enseñar la lengua japonesa a los extranjeros⁸. En esa muestra las clases de sílabas se distribuyeron de este modo: 5 del tipo V donde entraron las vocales; 4 VV formadas por los pares *ia*, *io*, *iu*, *ua* en símbolos fonéticos /a, o, u, wa/; 70 CV formadas

⁹ B. Bloch, *Studies on colloquial Japanese IV: phonemics. Language* 26: 86-125, 1950.

por 18 consonantes y una vocal y 21 CVV en las que 7 consonantes se combinaron con los pares *ia*, *io*, *iu*.

Los sonidos fueron grabados en la Broadcasting Science Research Laboratories (NHK) de Tokio. Con ese material se prepararon las pruebas que luego se realizaron en nuestro Laboratorio. Participaron diez oyentes argentinos, nativos de Buenos Aires.

De acuerdo a la instrucción que habían recibido cada participante fue transcribiendo las sílabas que oía utilizando los símbolos de su propio alfabeto. Ninguno de los oyentes tenía conocimiento del idioma japonés ni recibieron entrenamiento alguno previo a la prueba.

Los resultados que obtuvimos aparecen en las tablas 2, 3 y 4. En la tabla 2 vemos que cuatro de las cinco vocales se identificaron en su totalidad. Con la /u/ que dio el 80% esta vez la confusión fue distinta, un participante la transcribió como una *g* sola y otro como la sílaba *be*. Las sílabas VV tuvieron un alto porcentaje de inteligibilidad. Solo un oyente cambió *iu* por *ie*.

Tabla 2

Sílabas V	Sílabas VV
/i/ 100	/ia/ 100
/e/ 100	/io/ 100
/a/ 100	/iu/ 90
/o/ 100	/ua/ 100
/u/ 80	

Porcentaje de reconocimiento de sílabas japonesas compuestas de una y de dos vocales por parte de oyentes de habla española.

En la tabla 3, primera columna, vemos un porcentaje significativo de reconocimiento de los sonidos periódicos. Cuando hubo errores *m* se cambió por *n* y viceversa *ñ* con *m* y *n* y *l* con *r*. Con respecto a los pulsos, segunda columna, no hubo errores en *b d y*

Tabla 3

Sonidos periódicos	Pulsos	Sonidos no periódicos
/m/ 98	/b/ 100	/s/ 90
/n/ 92	/d/ 100	/x/ 97.5 (j)
/l/ 96	/g/ 98	/z/ 50 (*)
/r/ 90	/p/ 98	/ʃ/ 77.5 (*) (sh)
	/t/ 80	/ʒ/ 65 (y)
	/k/ 100	/tʃ/ 87.5 (ch)
		/ts/ 70 (*)
		/f/ 60

Porcentaje de reconocimiento de sonidos consonánticos japoneses por parte de oyentes de habla española. Los sonidos indicados con un asterisco no se usan en español.

/k/. En pocos casos hubo alguna confusión con *t* que se transcribió como /k/ y *p*. De todos modos, ninguno de los errores que señalamos en estas dos columnas tiene mayor peso ya que suelen producirse entre hablantes y oyentes de la misma lengua. En cambio en los sonidos que no tienen periodicidad registramos algunas diferencias. En la tabla 3 vemos que de los cinco que, en principio, serían comunes, tres fueron bien reconocidos *s* /tʃ/ /x/ y dos dieron porcentajes relativamente bajos /ʒ/ el 65% y *f* el 60%. Los oyentes pusieron también en símbolos los tres sonidos japoneses que no les eran conocidos /ts/ /ʃ/ y /z/. Al primero lo transcribieron como *t* y *s*, al segundo con la *y griega* y al tercero indistintamente con *y griega*, *ch* y *s*.

Según mostramos en la tabla 4, la sílaba CVV resultó en general menos identificada que la CV. Cuando hubo confusiones *l* se transcribió como *r*; la sílaba /kia/ se transformó en /tʃa/ sin vocal *i*. Algo parecido sucedió con *g* que se cambió por *d* en [dia] y por /ʒ/ en [ʒia]. Por último, el sonido /x/ se corrió principalmente a otro ruido fricativo /ʃ/⁸.

Tabla 4

Sonidos periódicos	Pulsos	Sonidos no periódicos
/m/ 86.6	/d/ 86.6	/x/ 56.6
/l/ 63.3	/g/ 43.3	
	/p/ 96.6	
	/k/ 70.0	

Porcentaje de reconocimiento de consonantes japonesas en sílabas CVV por parte de oyentes de habla española.

Resumiendo entonces encontramos que la mayor parte de las sílabas japonesas que presentamos a los oyentes españoles fueron bien identificadas. De las 100 sílabas que utilizamos 81 resultaron conocidas y de entre estas 55 alcanzaron el 100%, 11 dieron el 90%, 9 el 80% y 7 el 70%. De las restantes 19 sílabas, 9 tuvieron el 60%, 5 el 50%, 4 al 40% y una el 30%. Con respecto a los tipos silábicos, la combinación CV se reconoció el 89% de las veces; las cuatro clases VV el 97,5% y las de tipo CVV alcanzaron un 72%.

En la etapa siguiente de nuestro proyecto decidimos verificar si los oyentes japoneses reconocerían con la misma facilidad los sonidos de nuestra lengua. En el momento actual estamos desarrollando parte de esa tarea y podemos anticipar algunos de los resultados. El material que estamos utilizando incluye vocales, sílabas de tipo V, VV, CV, CVC, CVV y palabras formadas por dos sílabas CV (CVCV). Las listas se han presentado a participantes japoneses estudiantes graduados del Departamento de Ciencias de la Conducta de la Universidad de Hokkaido, Sapporo y del Departamento de Ingeniería Electrónica de la Universidad de Tokio. Se instruyó a los participantes para que luego de oír los sonidos los repitieran en voz alta para grabarlos y a continuación los representarían en símbolos. En una parte del trabajo se pidió que utilizaran caracteres japoneses y en otra que transcribieran en letras románicas.

Los resultados preliminares nos permiten anticipar que los oyentes japoneses tienen la misma facilidad que los españoles para

distinguir los sonidos de la otra lengua. Las vocales *i e o u* se reconocen directa y rápidamente, alcanzando casi el 100% de identificación tanto cuando se presentan en forma aislada como en contexto de sílaba y de palabra. Cuando se confunde la vocal *u* se cambia por *o*. Esto ocurre particularmente delante de algunas consonantes como *d*. Las sílabas VV, CV se identifican también con altos porcentajes. Estos son algo menores para las del tipo CVV y CVC. Por otra parte las palabras formadas por dos sílabas consonante-vocal (CVCV) como *poda*, "kapa", "kita", *dedo*, "dike" y *dote* alcanzan el 100% de reconocimiento.

Hasta aquí nos hemos ocupado de segmentos fonéticos tomándolos como eslabones separados de la cadena de habla. Sin embargo sabemos que en el habla espontánea los sonidos verbales fluyen en forma continua y se modifican por efecto del contexto. Por eso es necesario tener en cuenta otros segmentos más extensos que los abarcan como el acento, el ritmo y la entonación. En la próxima etapa es de rigor efectuar un cotejo entre las características suprasegmentales de las dos lenguas, en especial las que se relacionan con la duración de las estructuras rítmicas y los intervalos entre acentos.

Aunque el tema de esta comparación interlingüística está todavía en sus comienzos las perspectivas de progreso en esta materia son promisorias. Con la aplicación de técnicas de computación para el procesamiento digital de señales de habla, podrán inclusive desarrollarse sistemas de análisis, síntesis y de reconocimiento automático de secuencias acústicas de una y otra lengua.

Por otra parte, además de las contribuciones teóricas y técnicas que puedan transferirse a otras disciplinas relacionadas con las lenguas, el problema de la comparación entre el japonés y el español abre de por sí vastas perspectivas en el área de la comunicación por habla. Si tenemos en cuenta el considerable número de hablantes que pertenecen a uno y otro grupo lingüístico las implicancias de orden cultural caen de su peso.

Agradecimiento

Deseo expresar mi agradecimiento al Prof. Satoru T. Aiba y al Prof. H. Fujisaki por la valiosa colaboración que me brindaron du-

rante mi estadía en el Departamento de Ciencias de la Conducta de la Universidad de Hokkaido, Sapporo, y en el Departamento de Ingeniería Electrónica de la Universidad de Tokio, respectivamente.