

UNIVERSIDAD AUTÓNOMA DE MADRID

Facultad de Filosofía y Letras
Departamento de Filología Inglesa
Doctorado en Lengua, Literatura y Sociedad
Itinerario de Gramática General y Comparada

Aspectos cognitivos y formales de
las consonantes silábicas en inglés:
por qué se forman y cómo se forman.

TESIS DOCTORAL

por Javier Silió Martínez
dirigida por Carlos Piera Gil y Manuela Romano Mozo

Madrid, 2015

**There is something in staying close to men and women and
looking on them, and in the contact and odor of them,
that pleases the soul well,
All things please the soul, but these please the soul well.**

A mi familia, muy especialmente

Dedicatoria agradecida

Con deleite me regocijo por las buenas acciones de todos los seres

SHANTIDEVA

Una tesis es un trabajo científico que requiere observación y análisis, y también reflexión enjundiosa. Pero cuando trabajamos en el campo de la lingüística, como ciencia humana que es, se añade un interés especial por esas miles de formas de ser humano que las lenguas reflejan, todas iguales y todas distintas; este aspecto le da un lugar muy especial entre las disciplinas científicas.

El viaje exterior encontró su mejor reflejo en un profundo viaje interior, por la introspección que siempre hace falta para investigar y por la empatía con cada mujer y cada hombre que resuena en el lenguaje. Este trabajo ha sido para mí un viaje de madurez, no sólo intelectual sino también emocional. No es poca cosa.

Y ante todo, esta tesis es un homenaje a las personas que quiero: mi familia, mis amigos, mis maestros. Entre las líneas del análisis fonológico estáis todos presentes.

Gracias a todos los que habéis sido mis maestros, desde que empecé mis estudios en el colegio. En la Facultad de Filosofía y Letras de la

Universidad Autónoma de Madrid, a todos los que tomasteis parte en mis estudios de la licenciatura de Filología Inglesa, donde siempre me habéis hecho sentirme valorado y querido: gracias por la libertad que he respirado en este departamento, donde tanto he aprendido y donde me he formado como persona.

Por animarme a trabajar en la investigación en lingüística, gracias a Teresa Solías, Antonio Moreno, Elena de Miguel, Laura Hidalgo y Ana Ardid en la licenciatura; Amaya Mendikoetxea, Ignacio Bosque, Ricardo Mairal, Esther Torrego e Iggy Roca en el doctorado.

Mi recuerdo agradecido para Ken Hale por ese respeto inmenso y luminoso para con todos y con todo. Por el camino que abriste, *gogoan izango zaitugu betirako*.

Gracias a Águeda Sanz y a Lourdes Muñoz por vuestro cariño y vuestra ayuda constante desde las secretarías de Doctorado y de Filología Inglesa.

Mi agradecimiento a la Universidad Autónoma de Madrid por las Ayudas de Doctorado y a la Consejería de Educación de la Comunidad de Madrid por las becas de Formación de Personal Investigador, que fueron esenciales para hacer posible esta investigación.

Por contribuir a este trabajo, menciono a Teresa Balló y Clara Benito; Román García, Manolo Tirado y Leonor Langa; Ivar Zapata, Teresa Silió, María Arias, Ana Caballero, Mariluz, Sergio López². También al Centro Budista Tibetano Thubten Dhargye Ling. Y tantos más que

no cabéis en la lista, pero estáis. Agradezco a Tobias Scheer su página web, inestimable.

Gracias a Manuela Romano por recoger el testigo con la mejor de las disposiciones posibles, y por tu aprecio, tu entusiasmo y tu honestidad intelectual.

Y muy especialmente, gracias a Carlos Piera, director medular de este trabajo. Por tu inspiración, por tu discernimiento y por darme plena libertad en todo momento. Y por saber escuchar. Gracias por tu confianza y respeto, por tu enorme calidad humana, por llevar a tan altísimo nivel “ciencias” y “humanas”. *Jai Bhagwan*.

Gracias a mis maestros de Oriente. A Sigung Lai Chin Wah, Sigung Ho Fatt Nam y Sifu Wong Kiew Kit de Shaolin Wahnam. A Geshe Tsering Palden y Geshe Ngawang Losel y a los lamas de Ganden Shartse. Y a Amji Lobsang. Por darme claridad mental, capacidad intuitiva, mucha fuerza y buen ánimo para abordar este trabajo. Y sobre todo, por ser ejemplos vivientes de tan altas enseñanzas. *Jai Bhagwan*.

Un abrazo largo y lleno de calor para todos mis amigos, a lo largo y a lo ancho y a lo hondo del corazón. Siento no poder nombraros de uno en uno, porque sois muchos y el espacio apremia. Pero sabed que os tengo siempre en el pensamiento, yo con vosotros y vosotros conmigo. Gracias a la vida por uno de los regalos más grandes que pueden recibirse: ser rico en amigos.

Y siempre, siempre, siempre: mi familia. Por encima de todo, a mis padres, por haberme dado todo lo que tengo y todo lo que soy en esta vida, porque vais conmigo en todo momento y porque estáis en cada línea de este trabajo. Para mis hermanos queridos: Cristina (con Nicolás y Fernando), Fernando y Teresa, en los días buenos y en las horas tristes: hoy me complazco en compartir esta tesis con vosotros. Para mis tíos, abuelos y demás familia. Y para David, Sergio, Joakinn y Dani en las cuatro esquinas del sol.

You are all loved more than you will know.

Índice

Dedicatoria agradecida	v
Terminología y notación	xiii
Abreviaturas	xv
0. Introducción	1
1. Las CS inglesas y sus características	21
a. Qué consonantes pueden ser silábicas	22
b. Consonantes que son silábicas en inglés	25
c. Rasgos y elementos que caracterizan a las CS	30
d. Características propias de las CS inglesas	34
2. Por qué se forman las CS	39
a. Ventajas generales de las CS	39
b. La métrica del inglés como estrategia cognitiva	40
c. La ‘hipótesis optimista’	42
d. Argumentos ante la ‘hipótesis pesimista’	43
e. La alternancia entre schwa y CS	51

3. Cómo se forman las CS (I)	61
a. Cómo organizamos el espacio	63
b. Sensibilidad al peso en función de la transparencia	73
4. Cómo se forman las CS (II)	81
a. Estructura general de las CS	81
b. Sentido de la ramificación: la hipótesis TVR	88
5. Cómo se forman las CS (III)	99
a. Procesos de reducción parcial (1)	99
b. Reducción parcial (2): la coalescencia	110
c. Formación de CS por coalescencia	114
d. Observación final	124
6. Cómo se forman las CS (IV)	129
a. El locus de la CS	129
b. Coalescencia de rasgos	134
c. Las CS y la estructura de núcleo vacío	139
d. Las CS como ‘estructuras africanadas’	145
7. Cómo se forman las CS (y V)	155
a. La proyección de las CS	155
b. Implementación fonética	163
8. Relación inicio/CS: el plano segmental	169
a. La relación entre las CS y su inicio	169
b. Influencia de la sonoridad	171
c. Influencia del lugar de articulación	175
d. Asimilación inicio/CS: las antigeminoides	178

e. Las líquidas y el contexto izquierdo	184
f. La ‘distancia cero’	187
9. Relación inicio/CS: el plano superior	193
a. Más sobre el plano silábico: las CS/V__	194
b. El subplano métrico y la ‘captura de mora’	198
c. Captura de mora y CS	205
d. Cómo influyen los cúmulos consonánticos	210
e. Conclusión	215
10. El contexto derecho	219
a. Los contextos consonántico y absoluto	220
b. El contexto vocálico	223
c. Opciones del contexto vocálico	227
d. Condiciones morfofonológicas: reconocimiento	237
11. Las CS menos típicas	247
a. Las otras CS del inglés	247
b. Mecanismos de formación de las CS menos típicas	254
c. CS típicas y CS atípicas	263
12. Cómo se forman las CS en checo	271
a. Estructura de las CS checas	271
b. Dónde y cómo se forman las CS checas	275
c. Proyección hacia el plano métrico	282
d. Las CS checas en su contexto	284

Apéndice I: La Fonología de la Rección	293
Apéndice II: La derivación fonológica	307
Apéndice III: El acento en latín	317
Bibliografía	343

Terminología y notación

Los términos técnicos los marcaremos en negrita a medida que vayan apareciendo, y además los pondremos entre ‘comillas simples’ cuando sean términos introducidos por nosotros. Asimismo, hemos subrayado algunos conceptos clave, para resaltarlos y facilitar la lectura.

De acuerdo con Scheer (2003), representamos las consonantes **sonantes** con el símbolo ‘R’ y las **obstruyentes** con una ‘T’. Y siguiendo la notación del alfabeto fonético internacional (AFI), sus contrapartidas **silábicas** se marcan con un **diacrítico suscrito**, por lo que tendremos ‘Ṙ’ y ‘Ṛ’, respectivamente. De igual manera, /l̩, r̩, n̩/ son las variantes silábicas de /l, r, n/. Sólo la /ŋ/ se marca con el diacrítico en alto, /ŋ̥/, para mayor claridad gráfica.

Los **elementos fonológicos**, característicos de la Fonología de la Rección, se representan siempre con negrita. Por ejemplo, **R** es un elemento coronal, que se distingue de la R propia de las consonantes sonantes por estar resaltado en negrita.

Las variantes de las consonantes silábicas las anotamos según el criterio del *Everyman’s Pronouncing Dictionary*. De este modo, la /ə/ indica que no hay silabicidad (p. ej. 'melən, 'seldəm); la consonante sin más o con un diacrítico inferior (p. ej. 'preznt, 'lɪsn, 'lɪsŋ̩, 'pensl,

'kaʊnsl, 'kaʊnsl̩ə) indican que la consonante es silábica; y la '(ə)' indica que la silabicidad es opcional o que alternan las dos variantes entre sí¹, una silábica y la otra no-silábica (p. ej. 'leŋθ(ə)n, 'bʊz(ə)m, 'reŋn(ə)l). Las **variantes secundarias** (de uso menos frecuente) irán entre [corchetes] (p. ej. 'kɑ:bən [-bn], 'krɒmw(ə)l ['krʌm-, -wel]), como en el *Everyman's*.

La palabra **núcleo** la reservaremos para hablar del núcleo silábico (que en Fonología de la Rección es simplemente el núcleo, el nodo N), y llamaremos **cabeza** a otros núcleos. También reservaremos el término **ataque** (o *ictus* en inglés) para referirnos al golpe sonoro que hay al principio de una secuencia sonora como se utiliza en música. De igual manera, Asimismo, utilizaremos la palabra **cúmulo** (*cluster*) para referirnos a los agregados de segmentos.

Hablaremos de **unidades de tiempo** (*timing units*) con el sentido específico que se les da en Fonología Autosegmental, donde denotan cada uno de los nudos o vértebras del esqueleto autosegmental.

No hemos traducido los términos ingleses, por ser muy accesibles.

¹ Wells a su vez divide en dos la posibilidad intermedia: preferentemente con ə, o preferentemente como ʀ.

Abreviaturas

A-C:	Alargamiento compensatorio
C:	Consonante
Co:	Coda
CS:	Consonante Silábica
FR:	Fonología de la Rección
GA:	<i>General American</i>
In:	Inicio silábico
M:	Supermora
μ:	Mora
N:	Núcleo silábico
MEOSL:	<i>Middle English Open Syllable Lengthening</i>
PCO:	Principio del Contorno Obligatorio (OCP en inglés)
PSS:	Principio de Secuenciación Sonora
R:	Rima
R:	Consonante sonora
RLA:	Raíz Lingual Adelantada (ATR en inglés)
RP:	<i>Received Pronunciation</i>
σ:	Sílaba
SPE:	<i>Sound Pattern of English</i> , de Chomsky & Halle
T:	Consonante obstruyente
V:	Vocal

0 Introducción

“Where there is love there is no effort”

AMMA

Las consonantes silábicas (CS) rompen con la norma básica de que toda sílaba tenga una o más vocales. Este carácter excepcional es justamente el que atrae nuestro interés, convencidos como estamos de que todo lo que se sale de lo ordinario nos da una visión más amplia y profunda de la realidad. Como hace notar Chomsky, las excepciones son un motor esencial de la investigación científica:

(...) in fact the kinds of data that, say, physicists use are extremely exotic. If you took a videotape of things happening out the window, it would be of no interest to physical scientists. They are interested in what happens under the exotic conditions of highly contrived experiments, maybe something not even happening in nature. Chomsky (1999: 5).

Las CS nos sacan de lo automatizado. De no ser por lo mal que se habían comprendido hasta ahora (salvo las aportaciones recientes de la Fonología de la Rección), quizá nunca nos habríamos replanteado el entramado fonológico en su conjunto. El estudio de las CS comprende sílabas y pies, pero trasciende ese ámbito para tocar aspectos más generales y profundos de la teoría fonológica. Las CS nos movieron a buscar una sola visión general y unificada de lo que hasta ahora habían sido partes distintas del entramado fonológico. Pero antes de entrar de lleno en el análisis, vamos a dedicar esta introducción a

pasar revista a la metodología usada para la investigación y a los principios teóricos que hemos asumido en dicha tarea.

Metodología para el tratamiento de los datos

Para estudiar las CS sería deseable contar con ejemplos extraídos de la lengua viva, pero hay algunos inconvenientes. El primero es la poca fiabilidad de los hablantes para producir muestras, porque tienden a usar un registro más formal cuando se les entrevista para un estudio lingüístico. En sus estudios sobre el habla de la población neoyorquina, Labov¹ comprobó que el sesgo en las muestras requeridas (*elicited*) es inevitable. Los encuestados se esfuerzan por utilizar la lengua lo más “correctamente” posible, y evitan hablar como lo hacen habitualmente, por considerarlo demasiado coloquial (“*too slangy*”).

Pero en vista de que las CS son propias del habla informal o cotidiana, hemos buscado entre las canciones en inglés para tener muestras de este registro. Aun así, no hemos dejado de consultar con hablantes nativos para confirmar algunos detalles.

El segundo problema es la inestabilidad de los datos, que no permite basarse en la introspección exclusivamente. Ni siquiera los hablantes nativos, incluidos los lingüistas, consiguen ponerse de acuerdo sobre la incidencia de las CS. Así, Webster da *people* con schwa, pero Jones la descarta por completo². Y Toft (2002: 141) recoge dos resultados significativamente distintos para una misma población de hablantes, porque en uno de los análisis da *roughen* con /ŋ/ para un 31% de las

muestras (2002: 141, diagrama 40), pero en el siguiente análisis son el 47% de los casos (diagrama 41).

Ante la ausencia de una fuente segura de datos, recurrimos al ya clásico *Everyman's English Pronouncing Dictionary*, el diccionario de Daniel Jones que ha sido actualizado sucesivamente por Gimson y Ramsaran. Este diccionario cubre el inglés británico, sobre todo de **Received Pronunciation** (RP),³ que es el dialecto estándar del inglés.

El dialecto estándar del inglés americano se llama **General American** (GA). Para no dejarlo de lado, siendo el que mejor conocemos (y que además permite tratar la /ɹ/, porque es rótico), hemos consultado el diccionario enciclopédico de Webster. No está especializado en pronunciación, pero, como señala Arthur J. Bronstein (1990), falta un volumen de referencia en este dialecto del inglés⁴.

Con el fin de ampliar y mejorar el corpus de consonantes silábicas, hemos añadido la consulta de un diccionario inverso (Muthmann 2002).⁵ El criterio definitivo para la recogida de ejemplos es mantener un equilibrio entre las fuentes externas y la introspección, siempre atendiendo a los aspectos fónicos y no a los ortográficos.

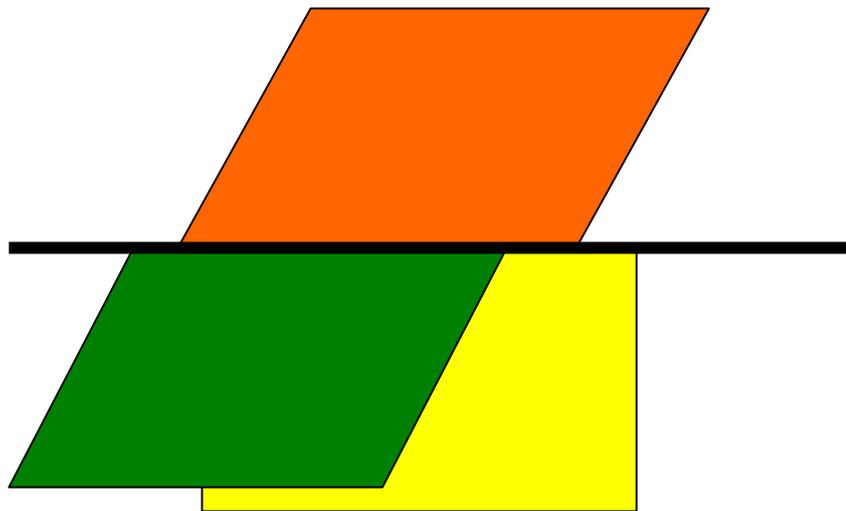
Ya que el análisis no podía ser del todo satisfactorio, al menos hemos querido contar con un número de ejemplos amplio, mayor de lo habitual en los estudios sobre CS. Los tratados fonológicos suelen limitarse a dar un puñado de ejemplos, y tienden a repetirse los mismos, sin prestar atención a los casos menos típicos. Nosotros vamos a tener en cuenta las CS en palabras monosilábicas, las CS en codas complejas, o la posibilidad de que las CS aparezcan en posición

media y no sólo en posición final. No dejaremos de lado las CS en derivados morfológicos, y atenderemos a la aparición de más de una CS en una misma palabra.

Marco teórico

Empezaremos por definir el **espacio fonológico**. Decimos que está formado por una serie de **planos** que se articulan entre sí, de acuerdo con las pautas de la **Fonología Autosegmental**. Cada aspecto de la información se organiza en un plano, y a su vez los distintos planos se articulan en torno a un eje o **esqueleto**:

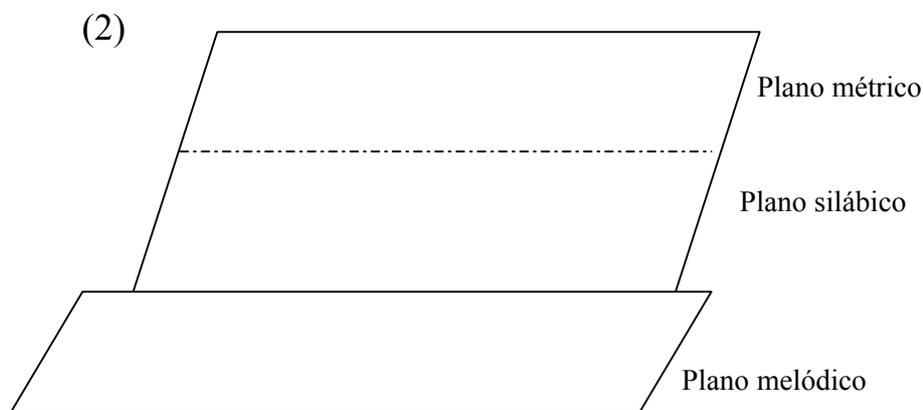
(1)



Empezamos por el **plano segmental** o **melódico** (*melodic tier*)⁶, que alberga las **unidades** fundamentales, los **segmentos fonológicos**. Estos se agrupan en unidades mayores, las **sílabas** y sus constituyentes, que están en un plano más alto, el **plano silábico**. Entre ambos planos media un **eje** que los articula y que sirve para

caracterizar el comportamiento independiente de cada nivel. A su vez las sílabas proyectan a un nivel más alto, el del **plano métrico**, formado por los **pies** y todas las proyecciones de la **jerarquía prosódica**.

Nosotros vamos a agrupar toda la información **suprasegmental** en un solo **plano prosódico**, que está formado por los **subplanos** silábico y métrico. Desarrollaremos esta idea en el cuerpo de la tesis, y por el momento nos limitamos a plasmarla en un gráfico:



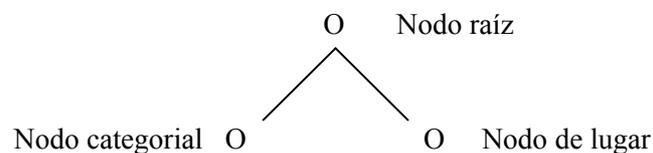
Nos quedamos ahora dentro del plano melódico para ver con más detalle los segmentos fonológicos. Los vamos a caracterizar por su **modo** y su **lugar** de articulación. El modo de articulación define la resistencia al paso del aire por el tracto sonoro, y sienta una división básica entre **consonantes** (C) y **vocales** (V), según haya o no un obstáculo en el flujo aéreo, respectivamente. Las consonantes más **resonantes** (*sonorant*) son las **sonantes**, clase que abarca las **líquidas** y las **nasales**. Menos sonoras son las **obstruyentes**, ya que ofrecen más resistencia al paso del aire, y en función de esta se subdividen en **fricativas** y **oclusivas**. Para facilitar el estudio, vamos a representar

las sonantes con una ‘R’, y las obstruyentes con una ‘T’, siguiendo la notación de Scheer (2003).

Las vocales y las sonantes son todas **sonoras**, porque interviene la vibración de las cuerdas vocales en su articulación, pero las consonantes obstruyentes pueden ser tanto sonoras como **insonoras**, esto es, sin voz.

Las características comunes permiten agrupar a los segmentos en **clases naturales**, cuyos integrantes compartan cierta propiedad. Y así podemos descomponer los segmentos en primitivos más pequeños. Cada segmento consta de un **nodo raíz**, del cual nacen el **nodo categorial** y el **nodo de lugar**:

(3)



Estos nodos llevan asociados a su vez toda una serie de **rasgos fonológicos** que definen el tipo de segmento⁷. De acuerdo con la **fonología no-lineal** o **autosegmental** posterior al **SPE**⁸, cada rasgo tiene su propia matriz, para señalar su autonomía.

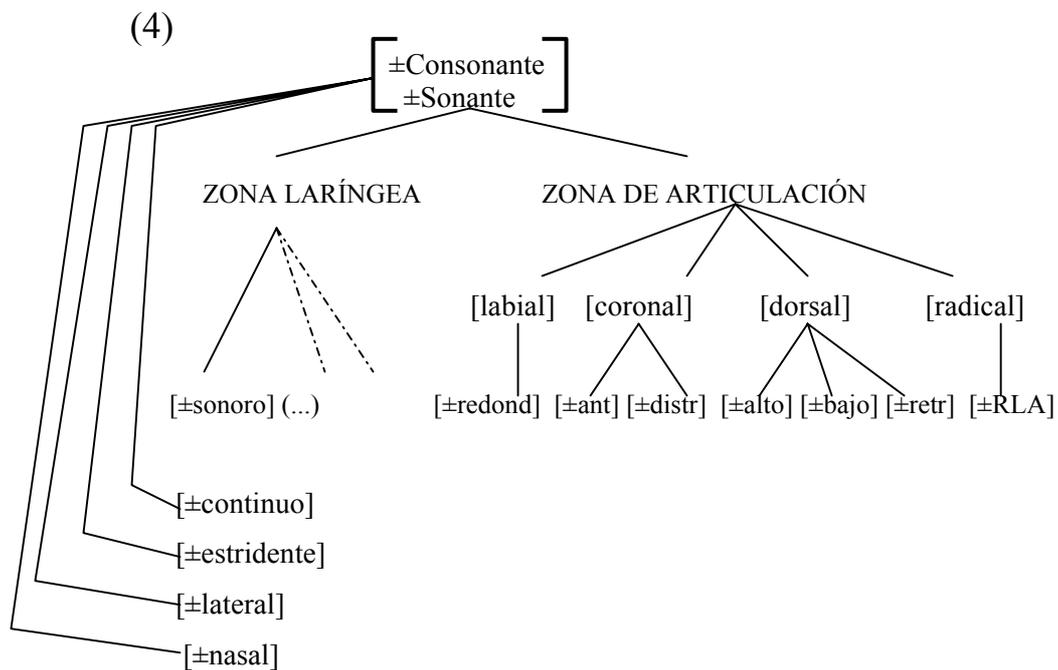
El nodo categorial define el modo de articulación. Esto se logra mediante los **rasgos de clase principales**, tales como el rasgo [\pm consonántico] (que separa a las vocales de las consonantes) y el rasgo [\pm sonante] (que distingue entre las consonantes sonantes y las

obstruyentes). Y en segundo lugar, mediante el **nodo de manera**, del cual dependen los rasgos [\pm continuo] (que diferencia oclusivas de fricativas) y [\pm sonoro] (para sonoros e insonoros).

Con respecto al lugar, la división primaria está en las cavidades faciales por las que sale el aire, que son la boca y la nariz. Así es como distinguimos entre segmentos **orales** y **nasales**, respectivamente, que se caracterizan mediante [\pm nasal]⁹. El lugar de articulación de las consonantes puede ser **faríngeo**, **velar**, **coronal** o **labial**. El lugar coronal¹⁰ puede ser [\pm anterior]: el **anterior** abarca la zona de los dientes y los alvéolos, mientras que el **posterior** corresponde al paladar anterior. El rasgo [\pm estridente] en su especificación positiva abarca /f, v/, pero en especial, las **sibilantes** /s, z, ʃ, ʒ/, que así se diferencian de /θ, ð/).

Las vocales se definen con el rasgo [\pm retraído] para los dos valores del eje horizontal (anteriores y posteriores), y los rasgos [\pm alto] y [\pm bajo] para tres grados de altura (altas, medias y bajas). La posición de los labios se describe mediante el rasgo [\pm redondeado]. Finalmente, la distinción entre vocales tensas y laxas (por ejemplo, /i-I/ o /u-U/) se realiza mediante el rasgo [\pm RLA].¹¹

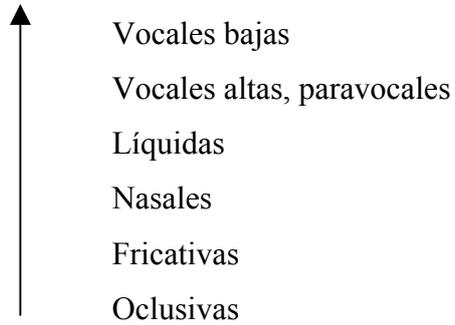
Todos los rasgos se estructuran en una **geometría de rasgos**, con una jerarquía y unas relaciones en concreto. Reproducimos aquí a modo de ejemplo el modelo que asume la *Nueva Gramática de la Lengua Española* (2011):



En la representación autosegmental, cada rasgo tiene su propia matriz, y los segmentos no son más que una coincidencia en el tiempo de determinados rasgos. Cada segmento se articula en torno a un nodo raíz, del cual dependen todos los demás nodos de rasgos.

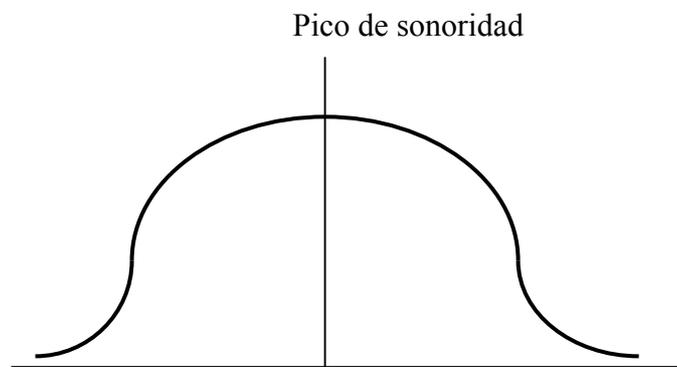
Dentro del plano melódico nos encontramos con secuencias de segmentos. No se distribuyen al azar, sino que se organizan en función de su sonoridad. Las **relaciones fonotácticas** determinan la capacidad de ciertos segmentos de estar uno junto a otro, y se establecen de acuerdo con la **jerarquía de sonoridad**:

(5)



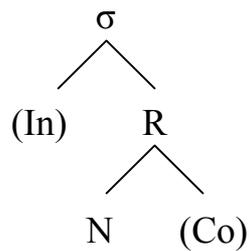
A lo largo de la cadena sonora, se repite un patrón que consta de un **pico de sonoridad** (correspondiente a una vocal) flanqueado a ambos lados por segmentos cada vez menos sonoros, tanto menos cuanto más lejos del pico de sonoridad. Este patrón se conoce como **Principio de Secuenciación Sonora (PSS)**.

(6)



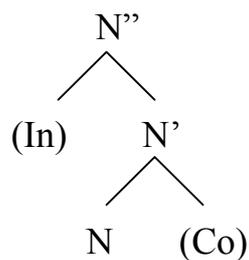
Esta información se refleja en el siguiente plano de la estructura, el plano silábico, que está formado por **nodos silábicos** abstractos:

(7)



El diagrama traduce a una estructura de constituyentes la idea de que la sílaba óptima consta de un pico de sonoridad que puede estar flanqueado por segmentos que van decreciendo en sonoridad a medida que se alejan de él. En concreto, la **sílaba** (σ) se organiza en torno a un **núcleo** (N), que es su **cabeza**, y que es el único elemento obligatorio. Este se asocia con la **coda** (Co) para formar la **rima** (R), y esta a su vez se asocia con el **inicio** (In). Coda e inicio son opcionales, y por eso las marcamos entre paréntesis. Es un formato basado en la teoría de la **X-barra**, con la rima como proyección intermedia (véase por ejemplo Rennison & Neubarth (2003):

(8)

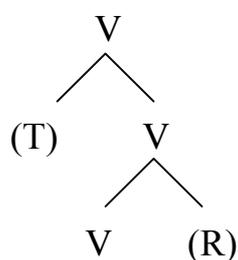


El esquema indica que la relación entre núcleo y coda es más estrecha, formando un **ámbito fonotáctico** (*phonotactic domain*), en tanto que inicio y núcleo gozan de independencia fonotáctica. Ewen & van der Hulst (2000) hacen notar que la relación fonotáctica sirve de argumento a favor de las sílabas.

También hemos querido explorar la representación de la estructura silábica con la **Estructura de Frase Escueta** (*Bare Phrase Structure*), que prescinde de los rótulos de nodo¹². Este formato no se ha utilizado en fonología, que sepamos, pero es posible que aporte buenas intuiciones en este campo, y por eso la hemos introducido aquí.

Aunque a lo largo de la tesis usaremos sobre todo la notación con nodos rotulados que es habitual en fonología, vamos a representar una sílaba mediante Estructura Escueta, para saber qué aspecto tiene: consta de un núcleo vocálico (V), una coda (con una sonante, R) y un inicio (que vamos a asignar a una consonante obstruyente, T):

(9)



Un aspecto de la Estructura Escueta que vamos a asumir es su carácter **derivacional**, que supone que la estructura se forma de abajo arriba. Desde luego, no optamos por un modelo **representacional** de estructura pre-formada en la que se inserten los constituyentes.

En la Fonología de la Rección se considera la posibilidad de que un **núcleo** esté **vacío**. Nosotros nos haremos eco de esta idea, si bien será en términos algo distintos de los de la FR¹³. Sea como sea, apreciamos

las intuiciones sobre las CS que debemos a esta teoría; en concreto, al trabajo de Harris, Scheer y Toft.

Para la asignación de la estructura silábica, se observa el **Principio de Maximización de Inicios**, que eleva la cantidad de segmentos iniciales de sílaba hasta el máximo que permitan las condiciones que cada lengua impone sobre la estructura silábica¹⁴.

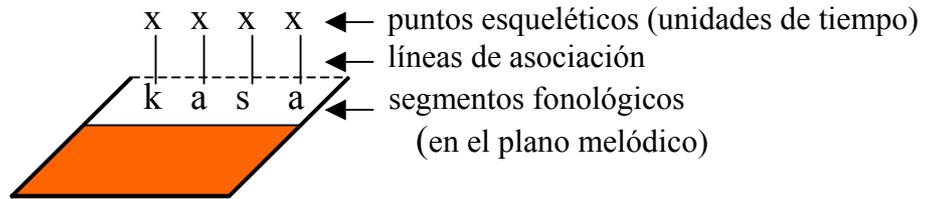
Los segmentos se proyectan desde el plano melódico a los nodos del plano silábico. Para conectar los dos planos, el segmental y el silábico, disponemos de un eje, al que se conoce como **esqueleto de tiempo** (*skeletal tier* o *timing tier*), una especie de columna vertebral formada por **unidades de tiempo** (*timing units*), que representan la **unidad de medida de tiempo fonológica** (*phonological timing*). Estas unidades de tiempo, sin contenido de rasgos, se llaman **puntos esqueléticos** (*skeletal points*) y se suelen indicar con una ‘x’:

(10)

x x x x x x x x x

La gradilla de tiempo capta el concepto de **cantidad fonológica**. Ya que el nivel melódico no usa información temporal (ni otros tipos de información relacional), las unidades melódicas se sincronizan con posiciones de medida temporal. La sincronización de segmentos y unidades de tiempo se indica con **líneas de asociación**, siguiendo el formalismo autosegmental:

(11)



En fonología autosegmental hay operaciones de disociación (*delinking*) y de expansión (*spreading*):¹⁵

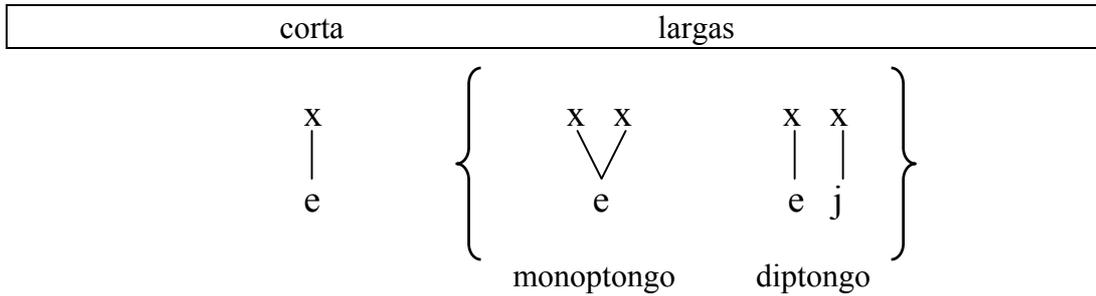
(12)



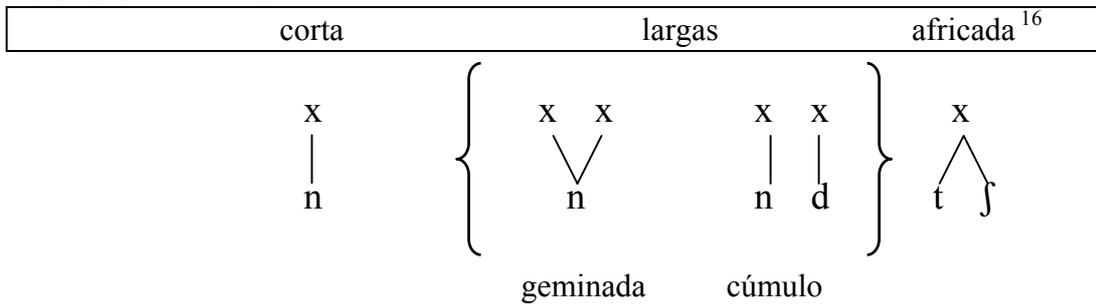
Este mismo formalismo autosegmental da una buena caracterización de la cantidad en vocales y en consonantes:

(13)

VOCALES:



CONSONANTES:



Así es como el esqueleto temporal articula segmentos y sílabas. Siguiendo el criterio del **peso por posición** de Hayes (1989), podemos determinar el **peso** de una sílaba por el número de elementos que tenga su rima. Las **sílabas ligeras** son aquellas que sólo tienen un elemento en la rima; dado que el núcleo es obligatorio, sólo cabe una vocal corta, y sin coda. Todas las demás son **sílabas pesadas**, porque cuentan con más de un elemento en la rima: puede ser una vocal larga o diptongo, y/o puede haber una o más consonantes en la coda. Las unidades de peso se llaman **moras**, y se representan con la letra ‘μ’; las sílabas ligeras tienen una sola mora, y las pesadas tienen dos.

El peso cuenta para el nivel siguiente, que es el nivel métrico, y que cuenta con su plano correspondiente. En el plano métrico se sitúan

todos los elementos de la jerarquía prosódica. En su nivel más básico, esta jerarquía cuenta con los **pies métricos**, que vamos a asumir aquí que tienen estructura binaria. Los **troqueos** tienen la cabeza o **prominencia** sonora a la izquierda, y los **yambos**, a la derecha:

$$(14) \quad \left[\begin{array}{cc} x & \\ x & x \end{array} \right] \quad \left[\begin{array}{cc} & x \\ x & x \end{array} \right]$$

troqueo yambo

Estamos utilizando aquí los diagramas de Hayes (1995), con paréntesis y aspas (*bracketed metrical grid*, que Hayes toma de Halle & Vergnaud 1987). También asumiremos su concepto de **euritmia** (1995: 372), que recoge la tendencia a que los acentos estén a una distancia ni muy larga ni muy corta, por lo cual tiende a evitarse el **choque de acentos**.

Los pies proyectan a niveles más altos, hasta llegar a la **palabra fonológica** y a la **frase fonológica**. No entraremos en esta cuestión, porque está fuera del alcance de esta tesis.

Eso sí, entenderemos que la estructura fonológica forma un todo, desde los rasgos infrasegmentales hasta los niveles más altos de la prosodia. En este sentido, podemos hablar de **proyección** de la información desde los niveles más bajos a los más altos.

La **arquitectura** o **estructura** se **construye** de abajo arriba, mediante **reglas**. Los planos delimitan ámbitos en los que operan las reglas,

pero en nuestro enfoque será posible que una misma regla traslade su acción a nuevos ámbitos para hacer al sistema más productivo.

La mayoría de las teorías al uso trabajan exclusivamente con una de dos posibilidades: o bien reglas o bien restricciones, pero nuestra postura es que el mecanismo fonológico tiene algo de las dos. El sistema nervioso central tiene este mecanismo de operación, y sospechamos que el lenguaje también se rige por este método. En concreto, pensamos que las **restricciones** son de salida (*output*), y vienen dadas en buena medida por los módulos o sistemas con los que interactúa el componente fonológico.

Desde un punto de vista más general, tomamos como objeto de estudio la **facultad del lenguaje**. Dado que el lenguaje (donde también incluimos a la fonología) forma parte del cerebro/mente de los seres humanos, nuestra perspectiva es la de la **biolingüística**, y muy concretamente, la de la **Biología Evolutiva del Desarrollo**, también conocida como **Evo-Devo**, por su nueva manera de concebir la evolución biológica y el diseño de los seres vivos. Huelga decir que en este enfoque es de esperar que muchas cualidades del lenguaje sean similares a las de los seres vivos.

Una idea fundamental y característica del enfoque evo-devo es que el diseño (y la evolución) de los seres vivos no suele hacerse sobre elementos aislados, sino sobre bloques enteros ver Sampedro (2002: 105). Por ejemplo, una misma estructura, toda entera, puede copiarse o repetirse en otro lugar, y de ahí que finalmente los seres vivos se parezcan tanto entre sí. Sampedro (2002) lo llama **evolución**

modular, y explica cómo difiere significativamente del darwinismo clásico:

Pese a las arraigadas convicciones darwinistas de mi juventud, he llegado a persuadirme, a base de palos propinados por la evidencia, de que (al menos en algunas de) las principales innovaciones biológicas de la historia de la Tierra tienen un mecanismo causal no darwinista, no explicable por la lenta acumulación de pequeñas mejoras adaptativas. He llegado a la convicción *científica* de que esas innovaciones tienen una naturaleza modular: consisten en la incorporación, o en la nueva utilización, de módulos genéticos completos y previamente funcionales. (2002: 137)

Sampedro da tres ejemplos concretos de evolución modular para fundamentar su postura. Uno es la simbiosis que dio lugar a las células eucariotas, abundantemente documentada por Margulis (véase Sampedro 2002: 31ss.). Otro es la duplicación de genes controladores de otros genes (la fila Hox). Como dice muy gráficamente este autor (2002: 141), “los genes *downstream* no se duplicaron físicamente, pero pasaron de tener un solo jefe a tener diez jefes en fila india”. Para entenderlo más a fondo, hay que ver que estos cambios pueden no generar estructuras nuevas, pero sí un potencial evolutivo nuevo (2002: 142).

El tercer mecanismo es la alteración del ADN regulador de un gen crítico, que se ve en los genes de polaridad segmental. Un detalle interesante de estos genes es que no están agrupados todos juntos, pero su actuación sí que ocurre de manera coordinada, en bloque, gracias a las proteínas que fabrican (2002: 143). Sampedro pone como ejemplo los círculos que se dibujan en las alas de las mariposas, y que no son otra cosa que una reutilización del módulo genético para fabricar metámeros (segmentos corporales). Se reutiliza todo un módulo completo, con todas sus propiedades integradas y coherentes. Como dice Sampedro, “el mismo gen también se expresa en otro

tiempo y lugar” (2002: 144). Esta idea de cómo se reutilizan mecanismos ya existentes, en conjunción con la importancia del lugar en el que operan, va a resultar fundamental para articular nuestro trabajo. Y desde luego, recomendamos muy vivamente la lectura del libro de Sampedro para entender todos estos mecanismos, porque estamos convencidos de que tendrán una enorme repercusión en la teoría lingüística.

En la teoría Evo Devo se dice que la evolución no modifica los cuerpos, sino sólo la receta para construirlos (los genes), y que el mecanismo es asombrosamente simple. No se producen nuevas creaciones desde la nada, a partir de cero. Se utilizan paquetes de información y se hacen nuevas combinaciones de lo que ya existe. Así se consigue obtener complejidad a partir de mecanismos sencillos. Esta idea nos llama la atención muy poderosamente, y además nos conecta con el **Programa Minimista**, aunque no entraremos en mayores detalles del mismo¹⁷.

Quedan así esbozadas las ideas de partida y nos ponemos en disposición de abordar el estudio de las consonantes silábicas del inglés.

¹ Véase Aitchison (2001: 43-49).

² También se aprecian discrepancias a la hora de establecer el silabeo en inglés. Incluso los expertos no pueden ponerse de acuerdo con una palabra tan sencilla y común como *happy*, según comenta Ladefoged (2001: 230).

³ La alternativa para el inglés británico es el diccionario de Wells.

⁴ El diccionario más utilizado para estudiar la pronunciación del inglés americano es del de Kenyon & Knott. Los atlas lingüísticos de Estados Unidos disponibles en internet no son de gran ayuda en el tema que nos ocupa, toda vez que la variación regional no afecta a las CS.

⁵ Al usar el diccionario inverso, hay que contar con que la ortografía no siempre orienta. Por ejemplo, hay que tener en cuenta *instinct*, *pharynx* o *peduncle* para estudiar la /ŋ/. También hay que contar con los gerundios <-ing> cuando tienen la pronunciación dialectal [-ɪn] o [-ən], que son los que informalmente se transcriben como <-in'>. De igual manera, se pasaría por alto que las palabras con <-ern> son candidatas a /ɪr/en los dialectos no-róticos.

⁶ El nombre deriva de los estudios no-lineales sobre lenguas tonales, donde hay melodías suprasegmentales de tonos, que son independientes de los aspectos no-tonales de una representación. En las lenguas que los tienen, los tonos constituyen otro plano o nivel que se liga al esqueleto autosegmental.

⁷ En otros sistemas, los segmentos se descomponen en **partículas** (Fonología de Partículas) o en **elementos** (en la FR). El sistema de elementos trata de reducir al mínimo el inventario de primitivos (que luego pueden combinarse entre sí), y son todos **monovalentes**: simplemente se tienen o no se tienen, pero no hay especificación positiva o negativa. (Ver Apéndice I, *infra*).

⁸ El *Sound Pattern of English* de Chomsky & Halle, que consignaba todos los rasgos en una sola matriz; esta era una **representación lineal**.

⁹ En algunos sistemas, el rasgo nasal es un rasgo privativo, que por tanto no tiene la especificación binaria (positiva o negativa) de [\pm nasal]. Nosotros nos limitamos a seguir el sistema de Ewen & van der Hulst (2000), sin entrar en este debate.

¹⁰ El término 'coronal' comprende aquí el ámbito más amplio que corresponde al nodo fonológico coronal, y no sólo a las articulaciones realizadas con la corona de la lengua, como aclara la *Nueva Gramática de la Lengua Española* (1.4e).

¹¹ Raíz lingual adelantada, o [\pm ATR] en inglés.

¹² Para una explicación en detalle de la Estructura de Frase Escueta, remitimos a Chomsky (1995), en Bošković & Lasnik (2007: 223-31) y a Lasnik & Uriagereka (2005: 31ss.).

¹³ Remitimos al lector al Apéndice I para una explicación más detallada.

¹⁴ Roca & Johnson (1999:279) da un principio de **Minimal Onset Satisfaction** que se enuncia como sigue: *Minimal satisfaction of onsets takes priority over satisfaction of codas*.

(1999: 283) expresa el sesgo a favor del inicio como **Onset Maximization**: *Maximal formation of onsets takes priority over formation of codas*.

Y finalmente, (1999: 284) agrupa los dos principios anteriores en uno más general, denominado **Onset First Principle**, que favorece los inicios por encima de las codas, y que a su vez puede derivarse de otro principio aún más general que favorece un salto máximo en la sonoridad al comienzo de la sílaba frente a una caída mínima al final (1999: 283).

Nosotros nos vamos a quedar con la idea de maximización de los inicios como forma simple de manejar todos estos principios.

¹⁵ Estas operaciones respetan el principio de no entrecruzar las líneas de asociación, igual que no se entrecruzan las líneas en los árboles de la sintaxis.

¹⁶ El **contorno** (otro término tomado del análisis no-lineal de tonos) capta la noción de que las africadas son cuantitativamente simples (ligadas a una sola unidad de tiempo) pero cualitativamente complejas (tienen dos unidades melódicas).

¹⁷ Véase Chomsky (1995), Radford (2004), Lasnik & Uriagereka (2005), Bošković & Lasnik (2007) para más detalles.

1 Las CS inglesas y sus características

Empezamos por acotar el campo de estudio. Las **Consonantes Silábicas** (CS) se han venido definiendo como “*consonante en posición de núcleo silábico*”, en una formulación que nos parece demasiado laxa, porque permite que cualquier consonante ocupe cualquier núcleo. Una sobregeneración tan fuerte no es deseable para una teoría, y como es de esperar no halla respaldo en los hechos observados. Es evidente que sólo a algunas consonantes se les permite ser silábicas¹, para empezar. Además, no todas las lenguas del mundo tienen CS: sobre una muestra de 182 lenguas representativas, Bell sólo atestigua 85 que tengan CS². Y aun dentro de las propias lenguas que las tienen, las CS tienden a ser poco frecuentes.

Vamos a matizar la clasificación tipológica, porque si bien las CS no abundan, tampoco son tan excepcionales. Y además nos llama la atención que las lenguas con CS son bien distintas entre sí, desde el salish al checo y el bereber, pasando por el inglés. Nos da la impresión de que ha habido una **convergencia evolutiva**, que se ha llegado a esta formación por vías independientes. Por eso es bien probable que las CS no hayan surgido de la nada, y caemos en la cuenta de que estas lenguas tan dispares tienen un factor común. Lo que vemos nosotros es que las CS aparecen en lenguas con un alto índice de **consonantización**, que admiten un gran número de consonantes en

sus sílabas, con gran libertad para formar cúmulos. Son las lenguas **CVC**, por oposición a las lenguas **CV**. Por esto mismo vemos difícil, si no imposible, la formación de CS en lenguas CV. Es decir, que no habrá CS si no se permiten consonantes en posición de coda (por ejemplo, en zulú), o incluso si sólo se impiden a final de palabra (en italiano o en telugu) o sólo en interior de palabra (en luo).³ En definitiva, cuanto más naturales resulten las consonantes en una lengua, más tendencia a la formación de CS debería tener.

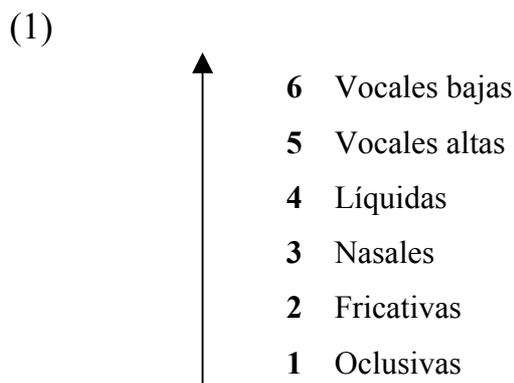
Pero volviendo sobre el carácter **marcado**⁴ de las CS, tiene que haber restricciones que limiten la presencia de consonantes en el núcleo de una sílaba. Por eso vamos a matizar la definición tradicional, y a tal fin comenzaremos por definir cuáles son exactamente las consonantes que pueden ser silábicas. Una vez sentada esta caracterización fundamental, el resto del trabajo de esta tesis estará dirigido a entender por qué se forman las CS y cómo lo hacen. En el porqué atenderemos a la diversidad de motivos que guían la aparición de CS, y ya veremos que no es sólo una cuestión de articulación y percepción. En lo que respecta al cómo, profundizaremos en los detalles estructurales y geométricos que hacen posible que una CS se licencie en una derivación fonológica.

Qué consonantes pueden ser silábicas

El **núcleo** marca la porción más resonante (*sonorant*) de la sílaba, el **pico sonoro**, y por ello lo natural es que esté ocupado por una vocal. Pero como dice Spencer (1994: 20), las CS “*pueden formar el pico de una sílaba y comportarse igual que las vocales*”. Colegimos así que

una consonante podrá formar un pico silábico en la medida en que se parezca a una vocal, es decir, cuanto más resonante sea. De acuerdo con este criterio de sonoridad, las mejores candidatas son las sonantes ('R'): primero las líquidas, y después las nasales. Las obstruyentes ('T') en cambio son raras como CS⁵, y dentro de esta clase las oclusivas resultan peores candidatas que las fricativas.

Con el fin de aquilatar aún más este criterio básico, proponemos que se aplique la **escala sonora** para delimitar **numéricamente** el repertorio de CS de una lengua. Recordemos que la jerarquía sonora tiene asignado un valor numérico para cada miembro de la escala:



A partir de estos valores numéricos de la escala sonora, empezamos por definir una diferencia de sonoridad o '**distancia sonora**', $\Delta(\text{SON})$, que separa a una candidata a CS de la vocal que ocupa el núcleo silábico⁶:

(2)

$$\Delta(\text{SON}) = (\text{SON}_V - \text{SON}_{\text{CS}})$$

tomando como índice sonoro de las vocales el más bajo de su rango, es decir, $\text{SON}_V = 5$. Gracias a la distancia sonora cuantificamos el repertorio de CS de las lenguas, que definiremos por *la diferencia*

máxima entre la sonoridad de las CS y la de las vocales que tolera cada lengua:

(3)

$$\Delta(\text{SON})_i \leq n$$

donde n es un número que se asigna paramétricamente para cada lengua, y que representa la diferencia sonora que no puede excederse. Nótese que la condición de distancia sonora se puede formular también de forma negativa, como $*\Delta(\text{SON})_i > n$.

Vamos a recorrer el espectro de las distancias sonoras posibles. Una diferencia $\Delta(\text{SON})_i$ de 1 (es decir, 5–4) significa entonces que sólo se permiten CS líquidas, como ocurre en “**eslavo**” (checo y eslovaco). Y una diferencia de 2 (o sea, 5–3) ya da cabida a las nasales, como encontramos en “**germánico**” (inglés y alemán). En la lengua “**bereber**” tashlhiyt imdlawn se da como posible cualquier tipo de consonante silábica⁷, por lo que la diferencia permitida llega hasta 4 (o sea, 5–1)⁸. Para las lenguas que no toleran CS, como el español o el francés, la diferencia de sonoridad es $\Delta(\text{SON})_i = 0$.

Vamos a ver algunos ejemplos de CS en distintos tipos de lenguas:

(4)

a. CS líquidas en “**eslavo**”:

en checo, *srdce* ‘corazón’, *krk* ‘cuello’; *vlk* ‘lobo’.

en eslovaco, *srna* ‘cierva’, *prst* ‘dedo’, *krč* ‘calambre’; *vlna* ‘ola’,
kľznut ‘deslizamiento’.

b. CS sonantes en “**germánico**”:

en alemán, *Wasserstoff* ‘hidrógeno’, *Sprudel* ‘agua mineral’,

*angeln*⁹ ‘pescar’; *raten* ‘aconsejar’, *bleiben* ‘quedarse’.

en inglés, *water*, *people*; *cotton*, *prism*.

- c. CS en **tashlhiyt imdlawn**: *fk* ‘dar’, *syf* ‘desvanecerse’, *tklf* ‘está sucio’, *tʃft* ‘aplastaste’, *sfqqst* ‘irritarle’, *tfktstt* ‘se lo diste’.

Gracias a los valores de la escala sonora hemos cuantificado los datos y hemos matizado más allá de la distinción básica entre sonantes y obstruyentes. Una ventaja añadida de esta formulación es que recoge también la **relación de implicación** de cada lengua con respecto a la sonoridad de sus CS, por la cual si una lengua tiene CS poco resonantes, entonces también las tiene más resonantes (pero no al contrario), es decir:

(5)

$$\mathbb{T} \supset \mathbb{R}$$

Consonantes que son silábicas en inglés

Hemos visto que el rango sonoro en inglés es $\Delta(\text{SON}) \leq 2$, de modo que las CS de esta lengua son sonantes (esto es, \mathbb{R}). Se habla de que se pueden oír también obstruyentes silábicas (\mathbb{T}) en formas rápidas o informales del discurso. Mora Bonilla (2003: 98)¹⁰ cita *s(u)ppport*, *s(u)ppose* o *f(o)rget*, por ejemplo. Y también el *Everyman's Dictionary* sugiere una CS en *sens(i)tive* y *diff(i)cult*¹¹. A nosotros nos queda la duda de hasta qué punto son verdaderamente silábicas, ya que Ladefoged (2001: 83) dice que “*A reduced vowel may be voiceless when it occurs after a voiceless stop (and before a voiceless stop)*”, y cita los cúmulos con /ɾ/ de **preparatory**, **spectrograph**,

introduction; y los cúmulos con /l/ de *replicate*, *complicate*. Sea como sea, las posibles obstruyentes silábicas son todas fricativas, por lo que el valor de $\Delta(\text{SON})$ podría llegar hasta el 3, pero nunca hasta el 4.

En vista de la duda en torno a las Ṭ , vamos a ceñirnos sólo a las sonantes (R), tanto líquidas como nasales. Hemos visto cómo la mayor resonancia de las líquidas las convierte en mejores candidatas a CS que las nasales¹². Pero a su vez cada clase admite matices, algo así como si los números de la escala sonora admitieran decimales. Dentro de la clase de las **líquidas**, creemos que la /r/ debe de ser más resonante que la /l/, puesto que es posible formar cúmulos /-r/ (en *gnarl*, *pearl* o *curl* y también en alemán *Karl*, o el dialectal *bitzerl* ‘un poco’, mientras que en sentido contrario estos cúmulos son impronunciables: *[-lr]. De ser cierta esta conjetura, la /r/ deberá ser algo mejor candidata a CS que la /l/.

A pesar de su mejor opción a ser silábica, el estatus de la /r/ como CS en inglés no se discute en los estudios especializados. Todo lo más, se da por sentado que es así, sin mayor explicación. Algunos relatos se apresuran a incluir la vocal /-ɜ:r/ dentro de la /r/ silábica, pero nosotros vemos argumentos que lo desaconsejan. Uno es que en inglés no hay CS en posición inicial absoluta de palabra, por lo cual *earl* jamás puede interpretarse como *[r]. El otro argumento en contra es que en una palabra como *bird* somos capaces de pronunciar una verdadera /r/ silábica [bɹd] que suene distinta de la vocal [bɜ:rd]. Además, en la pronunciación [bɜ:rd] se puede alargar la vocal (por ejemplo, en un uso expresivo o por razones métricas), mientras que si

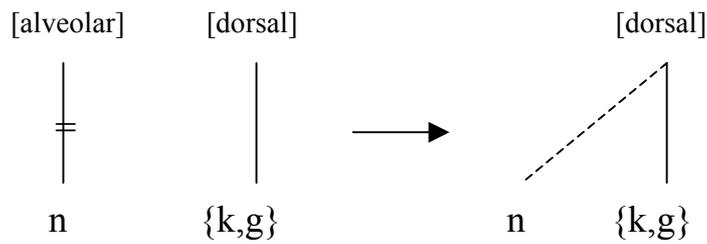
alargáramos la [r] en [bɹd], la /r/ tendería a sonar como una ‘vibrante múltiple’.

La /r/ que se hace silábica en inglés es la aproximante palatal /ɹ/, mientras que la del checo es una vibrante alveolar [r] (Spencer (1994: 20)). La gran resonancia de la [r] checa la capacita incluso para llevar el acento de palabra, mientras que su alófono palatal /ɹ/ (escrito <ř>), menos resonante, no es capaz de formar CS¹³, como demuestra Scheer (2009).

También dentro de la clase natural de las **nasales** se aprecian subdivisiones o matices. Spencer (1995: 20) constata que la más común como silábica es la /n/, seguida a distancia por la /m/, y con la /ŋ/ como opción prácticamente residual. Por lo tanto existe una correspondencia entre el grado de silabicidad de las nasales y su frecuencia de aparición en la posición de coda; es más, suponemos que hay una explicación fonológica basada en la asociación entre las codas y las R̥. Vamos a verlo.

En inglés la /ŋ/ es exclusiva de las codas. Prueba de ello es que nunca aparece a principio de palabra ni entre vocales. En lenguas como el vietnamita, tagalo o koriako, sin embargo, la /ŋ/ dispone de ambas posiciones silábicas, coda e inicio¹⁴. La diferencia fonológica la atribuimos a que la /ŋ/ del inglés es derivada: partimos de una /n/ subyacente, que en contacto con una oclusiva velar asimila su rasgo de lugar. La fonología autosegmental ha descrito muy adecuadamente este proceso como expansión (*spreading*) del rasgo [dorsal]:

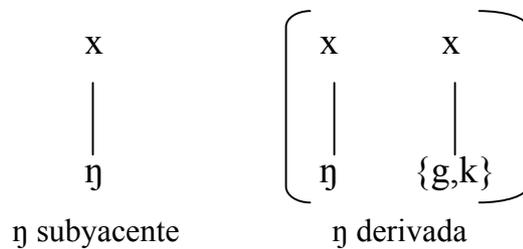
(6)



N.B.: La línea de asociación interrumpida entre la /n/ y el rasgo [alveolar] marca la disociación, a la cual sigue una reasociación entre la /n/ y el rasgo [dorsal], que se indica mediante la línea punteada o discontinua.

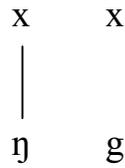
En cambio, la /ŋ/ de las otras lenguas que hemos citado es subyacente: no depende del contexto (recordemos que puede aparecer a principio de palabra o entre vocales) y no se ha formado por asimilación con otra consonante. Postulamos entonces que la η subyacente tiene una estructura simple, mientras que la η derivada tiene siempre una estructura compleja:

(7)



Es más, postulamos que en inglés la oclusiva velar puede quedar como elemento **flotante** (esto es, sin asociar), que en el curso de la derivación puede licenciarse (como en *finger* ['fɪŋgər]) o puede no hacerlo (es paradigmático el caso de *singer* /'sɪŋər/ o *singing* /'sɪŋɪŋ/, donde la /g/ no tiene nunca realización fonética):

(8)



estructura flotante

En cualquier caso, nuestra hipótesis es que la /ŋ/ del inglés (en contraste con las formas subyacentes simples que hemos visto en otras lenguas) es siempre un **segmento complejo**, y como tal se resiste a ser silábico.

Una dificultad añadida que suponen los segmentos complejos parece ser la resistencia (o imposibilidad, incluso) para que una vocal se reduzca a schwa en su vecindad. La schwa es necesaria en la formación de CS en inglés (véase *infra*), y en concreto la /ŋ/ debe formarse a partir de una /-əŋ/. Pero la realidad es que no hay ni un solo caso de /-əŋ/ en inglés, como hacen notar Ewen & van der Hulst¹⁵.

Nuestra hipótesis de que las estructuras complejas hacen más difícil la reducción a schwa encuentra apoyo en el *Webster*. Este diccionario tiene como norma reducir la vocal de las palabras terminadas en <-in> (por ejemplo, *robin*, *napkin*, *vermin* son todas /-ən/)¹⁶. Pero el criterio cambia cuando la secuencia <-in> está en un cúmulo, esto es <-inC> (por ejemplo, *hyacinth*, *labyrinth*, *absynth*); en estos casos la vocal se mantiene como /-ɪn/, sin reducir. Esta falta de reducción se aplica a las formas con <-ing>, tales como los gerundios con *-ing* y los diminutivos con *-ling*, así como alguna -ŋ/_k, (p. ej. *instinct* o *pharynx*). Podemos atribuir esta distinción a que el *Webster* está

considerando la nasal velar como una forma compleja (análogamente a las palabras que se escriben con <-inC>).

Después de todas estas consideraciones, sorprende encontrarnos con un tipo de /ŋ/ que sí que es silábica en inglés. Es la /ŋ/ que se forma por **asimilación progresiva**, en vez de **regresiva**, como sucede en *bacon* (*Everyman's*) o en el auxiliar *can* cuando es átono (“*I can try and book some seats round the corner*”, en Cruttenden (2001: 293))¹⁷. Nos parece lógico que este proceso opere con un sistema de reglas ordenadas. El primer paso es hacer silábica la nasal alveolar /n/, por el procedimiento habitual, y sólo entonces se asimila su lugar de articulación al de la oclusiva velar previa:

(9)

- 
- a. /n/ → /ŋ/ ('beikən → 'beikŋ)
b. /ŋ/ → /ŋ̥/ ('beikŋ → 'beikŋ̥)

O sea que la /ŋ/ ha cambiado su lugar de articulación, sencillamente. Desde esta perspectiva, la /ŋ/ por sí misma no es incompatible con la silabicidad, siempre que cumpla con la condición de ser un segmento simple. Si es así, la estructura juega un papel importante en la formación de CS.

Rasgos y elementos que caracterizan a las CS

La distancia sonora $\Delta(\text{SON})$ nos ha dado una idea muy intuitiva de las condiciones que una consonante debe cumplir para ser silábica.

Simplemente, hay una escala numérica lineal, y lo único que hay que hacer es ir dando pasos a lo largo de ella para aumentar o restringir la tolerancia de un pico silábico.

Ahora bien, la asociación entre consonante y núcleo silábico pertenece a la representación fonológica:

(10)



y aquí los segmentos no llevan asociado un número o índice que indique su grado de sonoridad, ya que están formados por rasgos (o por elementos, según la teoría que se utilice). Por eso tenemos que encontrar la manera de trasladar las aproximaciones de la escala sonora a la representación de los segmentos, que son los verdaderos candidatos a ser CS.

Las categorías **sonante** ('R') y **obstruyente** ('T') de la escala sonora que hemos estado manejando (y que es la más habitual) encuentran perfecto acomodo en los valores del rasgo fonológico [\pm sonante]: 'R' es [+sonante] y 'T' es [-sonante]. Por su parte, las subdivisiones de 'R' y 'T', que son líquidas y nasales, fricativas y oclusivas, se captan añadiendo el rasgo [\pm continuo]. Combinados ambos rasgos, se obtiene una caracterización de los cuatro tipos de consonantes de la escala sonora:

(11)

[+sonantes] [+continuas] = LÍQUIDAS (4)

[+sonantes] [-continuas] = NASALES (3)

[-sonantes] [+continuas] = FRICATIVAS (2)

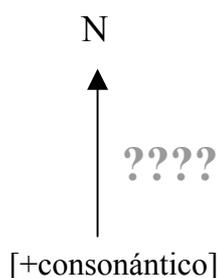
[-sonantes] [-continuas] = OCLUSIVAS (1)

N.B.: Hemos anotado el valor de la escala sonora junto a cada tipo de consonantes, a fin de facilitar la correlación con los rasgos.

Los rasgos que estamos manejando pertenecen a la clasificación de Ewen & van der Hulst (2000: 10-14), que subsume los rasgos [\pm sonante] y [\pm continuo] dentro de los llamados **rasgos de clase**; ¹⁸ esta categoría está dominada por [\pm consonántico], el rasgo que separa a las vocales de las consonantes. El dominio que [\pm consonántico] tiene sobre los demás rasgos de clase es significativo, en tanto que cada valor del rasgo principal lleva subordinado su propio conjunto de rasgos. Por ejemplo, el valor negativo [-consonántico] lleva asociados los rasgos [alto], [bajo], [retraído] y [tenso], mientras que el valor positivo, [+consonántico], se asocia a [continuo], [distribuido], [coronal] o [dorsal], entre otros.

Así es como el rasgo [\pm consonántico] nos puede dar la clave exacta del carácter marcado de las CS. Si lo normal (‘no-marcado’) es asociar el núcleo silábico ‘N’ al valor [-consonántico] propio de las vocales, la anomalía reside en asociar ‘N’ con un valor [+consonántico]. Así de concreto:

(12)



con todas las consecuencias que ello acarrea, por la diferencia en los rasgos que van subordinados a este cambio de valor.

Aquí surge una cuestión importante, además. La división cualitativa entre consonantes y vocales que los rasgos codifican, y en concreto a partir del rasgo fundamental [\pm consonántico], hace que el espectro de la escala sonora deje de ser continuo. Esta separación cualitativa entre vocales y consonantes tiene la ventaja de plasmar la tendencia general a que los núcleos silábicos estén ocupados por vocales¹⁹.

Por su parte la escala sonora tiene la ventaja, ya señalada, de indicar cómo el paso de vocales a CS es sólo una cuestión de grado²⁰. También recoge la relación de implicación, que además sirve para detallar qué tipos de consonantes son más aptas para ser núcleo de su sílaba.

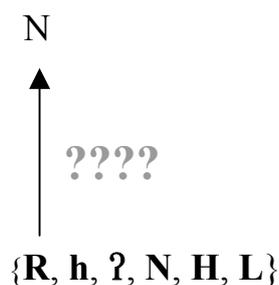
No queremos terminar este capítulo sin tomar en consideración el sistema de **elementos** unarios que la **Fonología de la Rección** (FR) utiliza, en lugar de los rasgos, para descomponer los segmentos. En desarrollos recientes de la teoría se han condensado los elementos en un repertorio único, que sirve para caracterizar tanto las vocales como las consonantes (muy especialmente, Harris 2004: 9-11). Esta falta de

distinción entre vocales y consonantes recuerda a la que ya hemos apuntado para el espectro de la escala sonora, y debe tener iguales ventajas e inconvenientes.

Al examinar en detalle la composición de los segmentos, sin embargo, comprobamos que sólo las consonantes están constituidas por todo tipo de elementos, tanto ‘vocálicos’ {**A**, **I**, **U**, **@**} como ‘consonánticos’ {**R**, **h**, **ʔ**, **N**, **H**, **L**}. En cambio los segmentos vocálicos se forman exclusivamente a partir de los elementos ‘vocálicos’ {**A**, **I**, **U**, **@**}. O sea que en la práctica se hacen distinciones dentro de los elementos básicos, y esa diferenciación afecta a la formación de CS.

Deducimos que cualquier asociación de un elemento ‘consonántico’ (o ‘no-vocálico’) con un núcleo silábico supone ya una anomalía, y se convierte en marca distintiva de las CS:

(13)



También da la impresión de que no todos los elementos son igual de válidos a la hora de asociarse a un núcleo. Hay una jerarquía donde el más apto es **R**²¹ y el segundo mejor es **N**, que son los elementos propios de líquidas y nasales, respectivamente. Aún más, el segmento más apto para formar una CS, que es la /r/, está constituido

únicamente por el elemento **R**. Y la /l/, el segundo más apto, también contiene **R** pero con la adición de **ʔ**. Con esta lógica se podría pensar que los segmentos con menos elementos son mejores a la hora de formar una CS.

Pero los datos de las consonantes nasales vienen a desmentir nuestro punto de vista, al menos según están especificadas en este momento. En concreto, la /ŋ/ es la más sencilla en su composición, únicamente **N** y **ʔ**, mientras que la /n/ y la /m/ constan además de los elementos **R** y **U**, respectivamente. Es posible que no haya una traslación lineal entre la composición elemental de los segmentos y su aptitud como CS; pero la teoría de los elementos no es definitiva y está en revisión permanente, por lo cual el estudio de las CS podría ayudar a refinarla. En cualquier caso, una mayor profundización al respecto queda fuera del alcance de esta tesis.

Características propias de las CS inglesas

Hemos visto a lo largo de este capítulo cuáles son las consonantes más aptas para convertirse en silábicas, y ha quedado claro que las reglas no son las mismas para todas las lenguas. No todas las CS son idénticas, hay variedad dentro de la categoría. Creemos que todas las CS del mundo deben compartir unos rasgos básicos que son los que les dan carta de naturaleza, pero al mismo tiempo se permiten diferencias entre lenguas a partir del modelo fundamental.

Es habitual que los estudios se centren en una lengua en concreto, y con ello se consiguen buenos resultados. Ahora queremos ir un paso más lejos; vamos a tratar de mejorar nuestras intuiciones sobre las CS inglesas (o ‘germánicas’) comparándolas con las de otras lenguas, y en particular las del checo (o ‘eslavo’).

La primera característica es el **tipo** de consonantes que pueden ser silábicas. Ya sabemos que en inglés son todas las R o **sonantes**, mientras que en checo sólo valen las líquidas y en tashlhiyt imdlawn se admite cualquier consonante.

La segunda es que las CS inglesas no se ven nunca en posición **tónica** o acentuada, a diferencia del checo y el tashlhiyt. La presencia en posición tónica es para Scheer (2009) una prueba de que las CS están asociadas a un núcleo silábico (a diferencia de las ‘consonantes atrapadas’, como veremos). Hay que decir que el carácter átono de las CS inglesas se debe a razones métricas, ya que aparecen siempre en la porción débil de un pie. Pero no por ello dejan de estar asociadas a un núcleo.

La tercera característica es que las CS inglesas pueden aparecer en contexto **vocálico**, CS/ __V, un hecho prohibido a los otros tipos.

La cuarta es que la CS casi siempre **alterna** con una schwa (por ejemplo, *bacon* puede ser tanto /'beɪkŋ/ como /'beɪkən/).

La quinta, en relación con la alternancia: las CS inglesas son **productivas**, pueden formarse CS en palabras de nuevo cuño, en tanto

que las del checo están congeladas léxicamente: hay un repertorio bastante fijo.

Una sexta condición es que las CS checas provienen tanto de secuencias TBRT como TRBT (véase Scheer 2009); es decir se desarrollaron diacrónicamente con la vocal fría (que era una ‘yer’) **antes o después** de la sonante R. Las R inglesas alternan exclusivamente con la vocal fría (que es una schwa) **antes** de la R (əR), desde el punto de vista sincrónico. Aunque es cierto que hay algunos casos formados (al menos en origen) a partir de Rə; esto lo veremos con más profundidad en el capítulo 11.

Sea como sea, un estudio de las CS debe tratar de explicar el cómo y el porqué de estas consonantes. Empezando por los rasgos generales que caracterizan a todas las CS respecto a otros segmentos, y ya dentro de las CS, dar cuenta también de las diferencias observadas entre las distintas lenguas; y en nuestro caso, de las características propias de las CS inglesas. Este es el objetivo que nos proponemos para todos los capítulos que siguen.

¹ Quizá la excepción sea el tashlhiyt, que al parecer permite que cualquier consonante sea silábica, como veremos seguidamente.

² En Greenberg (1978: 155). La cifra dada por Bell debe mejorarse al máximo, para ver exactamente cuáles son las lenguas que verdaderamente tienen CS.

Nosotros teníamos dudas sobre las CS en las lenguas Níger-Congo. Hayes cita nasales silábicas en gokana (1989: 356, basado en Hyman 1985) y en kimatuumbi (1989: 355, basado en Odden 1981), pero no dejaba de ser una referencia teórica. Sin embargo, nos hemos convencido de su existencia después de escuchar a una hablante nativa de suahili, Mary Kimai, de Tanzania. Se escuchaba una CS muy nítida tanto en *mtu* ‘hombre’ como en *ndiyo* ‘sí’. Es significativo que en *mtu* es la nasal silábica la que porta el acento de palabra.

³ Ver Harris (1994: 160-62).

⁴ Blevins (2004) hace una distinción general entre lo marcado y lo no-marcado o natural en los siguientes términos: *Marked patterns are rare, unnatural, and non-optimal, while unmarked patterns are thought to be frequent, natural, and optimal*. Para una explicación detallada de este concepto, véase (2004: 74-78).

⁵ Scheer (2003b) descarta directamente que las **obstruyentes** puedan tener estatus **silábico**.

Pero quizá deba admitirse que hay CS así en lenguas del Pacífico norteamericano, por mucho que el criterio haya sido poco restrictivo. En este sentido, nos parece elocuente y fiable el ejemplo tomado del Bella Coola *tlp^htyχ* ‘cortarlo’ que se cita en Toft (2002), tomado de Laver (1994). También Spencer (1995: 20) dice que *a number of native languages on the northwestern seaboard of America have such consonants*.

Nos queda claro que la sibilante [ʒ] de *vždy* en checo y eslovaco no es silábica, y más nos inclinamos por considerar que forme una africada inicial con la [v]. Ahora bien, en ruso, las preposiciones *v*, *k* y *s* están formadas por una sola consonante. Cuando se pronuncian aisladamente, parece que hay una schwa, pero cuando están en una frase (p. ej. *Ya poiedu v Moskvu*) no hay [ə] y la preposición suena como si fuese una coda. Especialmente convincente nos resulta el estudio de Yu (1999) sobre las sibilantes silábicas del mandarín. Y de paso nos sugiere que las sibilantes son, dentro de la clase de las fricativas, las más idóneas para convertirse en CS.

El bereber llama la atención por sus obstruyentes silábicas, tanto fricativas como oclusivas. Ridouane (2002) nos parece una fuente fiable en este sentido (véase *infra*). También Clements (en Roca 1997: 289-330) recoge la palabra (t̪f)(t̪k̪t̪) [la consonante subrayada es la silábica, y los paréntesis marcan lindes silábicas en la notación de Clements] en su comentario a Dell & El Medlaoui (1985).

⁶ Siguiendo con la escala numérica, al hablar del contexto izquierdo veremos cómo también se marca una distancia sonora entre la CS y el inicio que la precede.

⁷ Se ha discutido si las CS del tashlhiyt imdlawn son consonantes puras o si en realidad hay presencia de vocales. El estudio de campo de Ridouane (2002) sale al paso de estas dudas y concluye que las CS son realmente consonantes, especialmente en los dialectos del Anti-Atlas y Haha; es cierto que se observan trazas vocálicas en algunos hablantes del dialecto de Agadir, que se deben a la influencia de la lengua árabe, más prestigiosa.

Aprovechamos para aclarar aquí que la palabra *berber* o *bereber* resulta ofensiva para este pueblo, que prefiere llamarse a sí mismo *amazight*. Pero vamos a mantener el término *bereber* para facilitar la exposición, por ser el de uso habitual en nuestra lengua.

⁸ El Bella Coola también parece admitir CS las obstruyentes, como hemos señalado.

⁹ Nos referimos especialmente a una pronunciación dialectal de los Alpes austriacos, donde *angeln* ‘pescar’ es [ˈaŋln] [comunicación personal de Andreas Werle, que fue profesor del Goethe Institut de Madrid y es hablante nativo de ese dialecto alpino].

¹⁰ Hay toda una serie de CS en posición prevocálica y pretónica que cita Mora Bonilla (diagrama (1) (c)): *correct, terrific, believe, police, computer, contain, malaria, forget, suppose*. Matizamos que *forget* está considerada en un dialecto no-rótico, del sur de Inglaterra, que es el objeto de estudio de ese artículo.

Las ʈ inglesas son principalmente la /s/ y la /f/, que son continuas y además son **estridentes**. (También caben las fricativas sonoras, aunque sean menos idóneas. El diccionario de Daniel Jones parece admitir silabicidad en la elisión vocálica de *Wriothesley* /ˈrɪəθslɪ/ o /ˈrɪðslɪ/ y de *clavecin* /ˈklæv(ɪ)sn/.) El rasgo de continuidad les permite extenderse en el tiempo, una pronunciación a lo

largo de las dos ‘unidades de tiempo’ de la rima (o sea que absorben la ‘unidad de tiempo’ de la schwa), si bien nos inclinamos a pensar que en la segunda ‘unidad de tiempo’ hay una **schwa insonora**, que está incorporada como margen derecho de la fricativa. La estridencia les proporciona un alto nivel de ruido, que pudiera resultar un factor decisivo, al igual que en la /r/ del checo (como vibrante múltiple).

Por cierto, las T que pueden ser silábicas en inglés, es decir /s, f, ʃ/, son las mismas que pueden formar apéndices a principio de palabra, como en *spring* o *stay*, en *phthaleine* o *phthalic acid* (también *phthisis* (y *phthisic*) puede ser /fθaɪsɪs/ en pronunciación más ‘culto’ o cuidadosa, si bien la forma principal de pronunciarlo es /θaɪsɪs/; la información es del *Everyman’s Pronouncing Dictionary*), y como en *schmaltz*, *schmuck*, *schnitzel* o *schnorkel*. Desde el punto de vista de Scheer, esos apéndices son en realidad ‘consonantes atrapadas’ (este concepto lo vamos a desarrollar en el capítulo 4).

¹¹ Queda por dilucidar también el posible estatus silábico de las consonantes que hemos marcado en negrita en “**D***’you have to let it linger?*” (en la canción del grupo irlandés The Cranberries), y en “*I could/couldn’t ‘ve made it without you*”.

¹² Al hablar del contexto izquierdo, veremos cómo las líquidas son compatibles con un mayor rango de contextos por la izquierda que las nasales.

¹³ Lo que sí forma la <ɣ> son ‘consonantes atrapadas’, que como veremos en el capítulo 4, se distinguen de las silábicas porque no tienen peso métrico ni capacidad de portar acento (por ejemplo, en *hřmět* ‘tronar’ o *hřbet* ‘espalda’, el acento lo lleva la vocal). Todo esto queda perfectamente argumentado en Scheer (2009), un estudio que recomendamos vivamente.

¹⁴ Por ejemplo, en koriako encontramos *ηəvojja* o *ηəvotkən* junto a *tətəŋvoŋ* y *taŋvoŋə* en el paradigma flexivo del verbo *ηəvok* ‘empezar’ [véase Spencer (1996: 131)]. El koriako es una lengua de la familia chukoto-kamchatka que se habla en Koriakia, territorio de la mitad norte de la península de Kamchatka.

¹⁵ Ewen & van der Hulst (2000: 18, nota 16) comenta que en inglés la /ŋ/ se asocia exclusivamente a vocales cortas o [-tensas]. Delante de ella pueden estar {æ, ε, ɪ, ɒ, ʌ}, y cabe incluso la /ʊ/ en un préstamo como *Jung*, pero nunca la /ə/ (*It fails to occur before /ŋ/*).

¹⁶ A principio de palabra se mantiene la /t/, sin reducirse a schwa. En casos como *insert* puede atribuirse a una influencia morfológica, pero no así en *insectivore*.

¹⁷ El contexto aún es más favorable cuando además sigue otra oclusiva velar, como en “*You can keep it in the fridge*” o “*You can go now*”, o también “*bacon crisps*” o “*bacon glamour*”. Por cierto, Daniel Jones no da [ŋ] en *deacon* ni en *beacon*, por lo que la frecuencia de uso de las palabras debe ser significativa en este caso, igual que en otros que comentamos más adelante.

¹⁸ Por su parte, los rasgos de lugar tienen menos incidencia sobre la sonoridad, como sabemos.

¹⁹ Al hilo de esta reflexión se nos ocurre una definición en negativo de las CS que puede resultar muy gráfica: *hablar de CS es hablar de sílabas sin vocal*.

²⁰ Esta misma continuidad que tiene la escala sonora, que es una medición acústica, existe seguramente en el aspecto articulatorio. Como señala Ladefoged (2001: 260), *This grouping reflects the fact that the four features, Stop, Fricative, Approximant, and Vowel, are all dependent on the degree of closure of the articulators. In some older feature systems, these possibilities are split into two groups, but it is now thought better to recognize that they form a continuum (...)*.

Ladefoged lo ilustra con un par de ejemplos concretos. Uno más concreto en danés, de donde proviene su apellido: *The stops first became fricatives, which later became approximants in Danish* [læðəfoɣəd] > [læðəfoɣəd], *making it apparent that there is a continuum going from*

[stop] through [fricative] to [approximant]. El otro ejemplo, más general, proviene del español: Spanish also has a process whereby stops first become fricatives and then approximants.

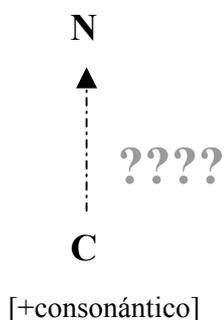
²¹ No se debe confundir el elemento unario **R**, que marcamos en negrita, con la abreviatura R que representa a las consonantes sonantes.

2 Por qué se forman las CS

Ventajas generales de las CS

Al inicio de este estudio hemos visto cómo las CS tienen un carácter marcado, que restringe mucho su presencia en las lenguas del mundo. Asociar una consonante (o un rasgo [+consonántico]) a un núcleo silábico ‘N’ tiene algo de anómalo:

(1)



El carácter marcado de las CS es un reflejo del esfuerzo que suponen para la economía lingüística. Las CS son un “lujo” que el sistema se permite, y hay que “pagarlo”: el esfuerzo debe compensarse con una ventaja en algún otro aspecto. Es decir, que si el sistema se presta a relajar o contravenir sus principios (estructurales o de otro tipo), algo bueno tiene que haber en las CS que justifique su existencia.

Queda claro entonces que las consonantes silábicas se forman para aportar un beneficio en algún ámbito o **módulo** de nuestro aparato lingüístico. Es bastante probable que esa ventaja no sea la misma en todas las lenguas; en concreto, sospechamos que en tashlhiyt las CS tienen utilidad fisiológica: gracias a ellas se consigue abrir menos la boca para hablar, y con ello se evita la deshidratación en un clima hostil¹. Para las CS del checo y otras lenguas eslavas no hemos encontrado mejor explicación, modestamente, que motivos estilísticos o subjetivos (alguna suerte de *Volksgeist*)². En cuanto al inglés, las CS se encuadran dentro del patrón métrico que esta lengua utiliza, y que entendemos como una estrategia cognitiva, destinada a facilitar el procesamiento de las unidades lingüísticas. Para entender mejor esta contribución, pasamos a revisar los fundamentos de esta estrategia.

La métrica del inglés como estrategia cognitiva

El ritmo de la lengua inglesa sigue una alternancia entre partes fuertes y débiles (ver, entre otros, Cruttenden 2001:250-55, Roca & Johnson 1999: 312-15, Spencer 1996: 282ss.). Este patrón facilita la articulación y la percepción, pero su utilidad va más allá de lo fonético y beneficia también a la economía cognitiva. La alternancia fuerte/débil resalta los componentes de más peso frente a los de menos peso, y con ello se mejora el procesamiento mental. La estrategia consiste en neutralizar o reducir los elementos **débiles**, que se difuminan y pasan a un segundo plano, donde quedan en semipenumbra; basta entonces con dejar los elementos **fuertes** como estaban para que por sí solos aparezcan resaltados (*salient*), en primer plano:

(2)

A a **A** a **A** a **A** a **A** a

Este contraste entre **primer plano** y **segundo plano** (*foreground* y *background*) permite un procesamiento más económico y eficaz de la información³. La producción, percepción y reconocimiento de las unidades lingüísticas se benefician de esta forma de organizar los elementos. En el momento de procesar la información somos capaces de restringir nuestra atención a los componentes del primer plano, que son sólo unos cuantos respecto al conjunto total. Y al mismo tiempo no dejamos de reconocer ese conjunto al completo, que siempre queda salvaguardado porque el segundo plano retiene la información suficiente para reconstruir el total.

La alternancia fuerte/débil en inglés produce un patrón métrico característico, que viene marcado por la isocronía entre sílabas tónicas. Aunque puede haber más de una sílaba átona en la parte débil, el patrón más habitual es una sucesión de **troqueos**, $(X x)_n$.⁴ A la hora de procesar la información, el primer plano de la atención va saltando de un acento a otro, mientras las sílabas de la parte débil se mantienen en la semipenumbra del segundo plano. A tal fin es característico que las vocales de la sílaba débil o átona se reduzcan. Las vocales altas pueden reducirse a /ɪ/ y a /ʊ/, pero la forma más general de reducción es a schwa, /ə/.

La ‘hipótesis optimista’

La hipótesis que planteamos es que las CS del inglés representan un caso extremo de este procedimiento. Asumimos que las CS quedan más en segundo plano todavía: reciben un peso cognitivo aún menor con el fin de facilitar aún más el procesamiento. El material que pasa a segundo plano en las CS es tan reconocible y permite reconstruir el original tan fácilmente, que resulta muy económico. Este esquema cognitivo cuadra con el dato empírico de que las CS tienden a aparecer en palabras de uso común; a nuestro entender, es en ellas donde se pueden limar más partes sin que se afecte demasiado el total.

Y por eso hemos dicho que en realidad las CS son un lujo que el sistema está en condiciones de permitirse. La idea es que las CS favorecen la estrategia cognitiva del inglés, pero sin que por ello se dañen dos puntos sensibles: el **reconocimiento cognitivo de las formas** afectadas, por una parte, y las **reglas de buena formación estructural**, por otra. Queremos poner nombre a esta propuesta, porque es la que da cuerpo a toda esta tesis: la vamos a llamar ‘**Hipótesis Optimista**’. Además de su ‘optimismo’ nuestra hipótesis es **holística**, porque tiene en cuenta todos los niveles de la estructura fonológica. Como iremos viendo a lo largo de la tesis, el patrón fuerte/débil del plano métrico afecta a todos los componentes de la estructura fonológica y llega hasta los mismos cimientos, hasta los rasgos infrasegmentales.

Nuestra hipótesis se desmarca así de la ‘**hipótesis pesimista**’ que se ha sostenido en la **Fonología de la Rección (FR)**. Como vamos a ver

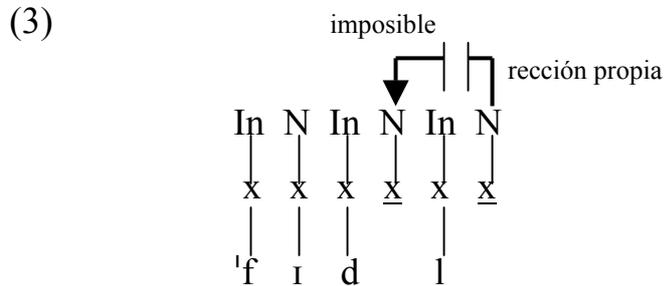
en la siguiente sección, la FR ha considerado a las CS como un mecanismo reparador para salvar una estructura mal formada, que si no sería incapaz de licenciarse o legitimarse. Esta es la esencia del debate. En tanto que la hipótesis pesimista atribuye las CS a un *por qué* (una ‘deficiencia’ estructural), en la hipótesis optimista las CS tienen un *para qué* (la ventaja cognitiva y métrica que justifica el precio que se paga por ellas); lo cual requiere de un *cómo* que permita implementarlo. Y a mayor abundamiento, también un *dónde*, porque cada hipótesis lleva aparejado su propio ámbito de extensión: la pesimista es muy **local**, porque se limita al plano segmental, en tanto que la optimista es más holística, como hemos dicho.

Argumentos ante la ‘hipótesis pesimista’

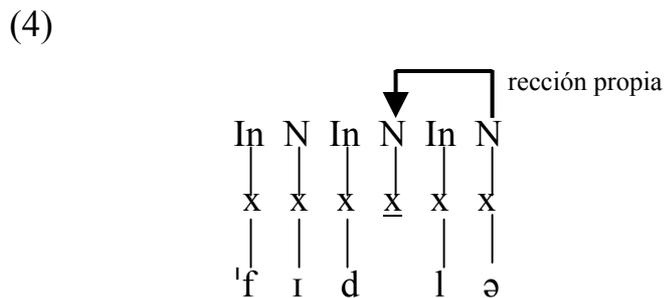
La teoría fonológica clásica habla de una secuencia mal formada que necesita ser reparada. En concreto, Roca (1994: 245) explica cómo al formarse una secuencia *-TR se rompe el contorno sonoro marcado por el PSS (Principio de Secuenciación Sonora), puesto que la consonante final es más resonante que la que la precede. Una secuencia así es impronunciable y reclama una solución, bien insertando una schwa (esto es, -TəR) o bien formando una CS (esto es, -TR̩).

La FR ‘clásica’ ha planteado el problema como una cuestión de **rección propia**⁵. Harris (1994: 184) analiza las sínopes como un término medio en la oscilación entre schwa y ‘cero’. Él parte de la base de que en *fiddle* hay un **núcleo vacío** interno, pero no puede

licenciarse porque no puede recibir rección propia del núcleo siguiente, ya que este también está vacío (1994: 192):

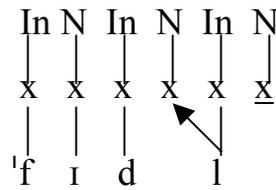


El contraste lo da la palabra derivada *fiddler*, donde sí que se puede licenciar el núcleo vacío interno (/ˈfɪdələ/), gracias a que recibe rección propia del núcleo lleno que le sigue:



Para solucionar la falta de rección propia en *fiddle*, según Harris (1994: 192), hay que dar expresión fonética al núcleo vacío, bien mediante la vocal fría latente (que en inglés es una schwa), [ˈfɪdəl], o bien como CS, [ˈfɪdɪ].⁶ También Toft (2002: 133-38) habla de formar una CS, donde la sonante (R) se expande hacia el núcleo vacío previo y así lo legitima:

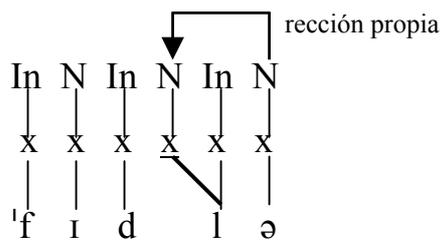
(5)



Durante años, Tobias Scheer ha sido un caso aparte dentro de la FR, al considerar que la secuencia -TR era un eslabón de un proceso de reducción vocálica progresiva. Pero Scheer (2009) se ha sumado a las demás voces de la FR, tanto Harris y Toft como Szigetvári & Blaho (2004)⁷, y ha pasado a defender la idea de que el núcleo de la CS está vacío subyacentemente.

Sin embargo, hay una batería de argumentos que a nuestro entender se oponen al criterio pesimista y local que se ha venido sosteniendo en la FR. El primero y más fundamental es que en inglés (aunque no en checo) hay CS seguidas por una vocal (el propio *fiddler* admite una CS, y no es un caso aislado, porque los ejemplos de CS/ __V en inglés abundan)⁸. Este argumento nos parece definitivo por sí solo para invalidar todo el razonamiento de la legitimación que ha aducido la FR; por mor de las CS/ __V, las CS dejan de ser necesarias para salvar una derivación.

(6)



El segundo argumento es que las CS en inglés son opcionales, como señala con mucho acierto Mora Bonilla (2003). Es conocido que en inglés la mayoría de las formas con CS alternan con una schwa⁹, que casi siempre hay una [-əR] que se corresponde con la [-ɹ]. Por ejemplo, *fiddle* puede ser tanto /'fɪdl/ como /'fɪdəl/, y *cotton* puede ser /'kɒtŋ/ o /'kɒtən/. Dado que el sistema dispone ya de la forma con schwa (-TəR) para reparar la secuencia anómala *-TR, la forma con CS (-TR) no parece que sea necesaria en absoluto.

La alternancia de formas nos sugiere de paso lo incierto que resultaría crear un mecanismo especial solamente para salvar una derivación perdida. Aunque no existiera la forma con schwa, que basta por sí sola para reparar la secuencia *-TR, no es muy económico formar CS *ad hoc* para solucionar el problema. A mayor abundamiento, el inglés dispone de un tercer mecanismo reparador, consistente en eliminar una de las consonantes que producen el conflicto, como se hizo con *sign* o *autumn*.¹⁰

Un tercer argumento es que las CS no sólo se forman en la base de la derivación fonológica, un aspecto que los estudios sobre CS han omitido sistemáticamente. Por ejemplo, dentro de la fracción léxica de la derivación pueden formarse CS tras la adición de **sufijos**, como sucede en los flexivos de *chosen, frozen, forgotten, forgettin', fadin'*¹¹ o en los derivativos de *wooden, harden, officer, official, survival*. Nos resistimos a creer que estos sufijos tengan en su forma subyacente un núcleo vacío, porque equivaldría a postular que el sufijo está formado únicamente por una consonante¹².

Ya en la fracción **postléxica**, vemos CS en concomitancia con la formación de ciertos **grupos clíticos**. No sólo ocurre con los verbos auxiliares contraídos (*it'll* o *that'll* con el auxiliar de futuro, y también auxiliares con negación como *doesn't*, *wouldn't* o *haven't*), sino también con otras categorías léxicas (*let'em*; *rock'n'roll*, *R'n'B*). Quizá se aprecie mejor la silabicidad de la CS de grupos clíticos en el contexto de una expresión completa¹³:

(7)

That'll be alright as it is.
You can do it better than me.
Take em in your hands.
I'd like some bread'n'butter, please.

Cuesta pensar que los monosílabos que dan lugar a la CS en estos casos estén vacíos de base, subyacentemente. Además, insistimos en que estas CS tienen que formarse con la estructura de pies ya asignada (en el plano más alto de la estructura fonológica, y no en el más bajo). Por eso pensamos que estas CS refuerzan la hipótesis optimista y holística, que va a asumir que las CS inglesas se forman en el ámbito de un troqueo.

Existe aún otra forma postléxica en relación con la parte métrica. Nos referimos a la formación de CS que sigue a un cambio en la ubicación del acento después de aplicar la **regla de desplazamiento del acento principal** en inglés (conocida como *English stress shift rule*: véase Hayes 1984 para una explicación detallada). Esta regla hace que el acento se retraiga para evitar un choque de acentos dentro de un sintagma nominal, y así es como una palabra (p.ej. *hotel*) con acentuación aguda (oxítónica) cambia su patrón a llano (paroxítono):¹⁴

(8)

$$\begin{array}{ccc} \begin{pmatrix} & x \\ x & x \end{pmatrix} \begin{pmatrix} x \\ x \end{pmatrix} & \longrightarrow & \begin{pmatrix} x \\ x \end{pmatrix} \begin{pmatrix} x \\ x \end{pmatrix} \\ \text{hoTEL} \quad \text{MANager} & & \text{HOtel} \quad \text{MANager} \end{array}$$

Al pasar del contorno **oxítono** subyacente al contorno **paroxítono** de la forma derivada, surge un contexto perfecto para formar una CS. Esta circunstancia nos reafirma en la hipótesis de que las CS no tienen un núcleo vacío subyacente, y de que en inglés las CS se forman en niveles altos de la derivación, ya en el ámbito de un pie.

Los argumentos finales, cuarto y quinto, vienen a sugerir que en la formación de CS no sólo intervienen condiciones estructurales, sean de rección o de otro tipo. El cuarto argumento es la relación entre frecuencia de uso y tendencia a formar CS. En el marco de nuestra hipótesis, las palabras de uso más frecuente son las más propensas a tener una CS. La razón es que es en estas palabras donde más se busca la facilidad de procesamiento cognitivo, y que en ellas podemos suprimir material fonológico sin que se afecte demasiado al reconocimiento de la palabra.

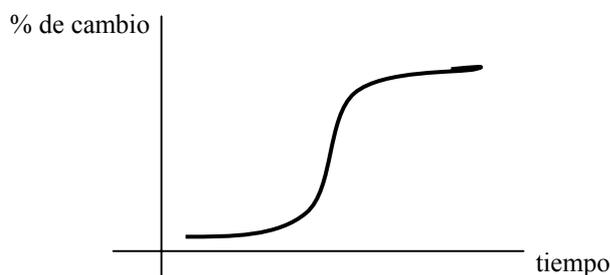
Por el contrario, en las palabras infrecuentes no se obtiene gran beneficio con la reducción fónica, precisamente por lo escaso de su frecuencia de aparición, y en cambio es fundamental conservar todo el material fónico justamente para preservar la información. De ahí que en estas palabras tienda a aparecer al menos una schwa, e incluso se conserve entera la vocal periférica cuando son términos científicos o técnicos. Por ejemplo, *anacoluthon*, *photon*, *neutron*, *argon*, *icon*,

benzol, *naphthol*, *argol* son todas con /ɒ/. Igualmente, *accent* y *Serpent* (la constelación) son con /e/, y *Senegal* es con /ɔ/.

Vamos a ver un caso que ilustra la relación entre frecuencia de uso y cantidad de material fonológico. Con idéntico contexto fonológico #v_r#, encontramos una **supresión** o ‘cero’ para la ubicua *every* y todos sus derivados (*-body*, *-one*, *-thing*, *-day*, *-where*, *Everyman*), que sin embargo alterna con schwa (es una /(ə)R/: véase *infra* la ‘silabicidad débil’) para las palabras de uso medio: *several*, *maverick*, *ivory*, *savoury* (y también *Savoury*), *favourable*, *favourite*, *flavouring*, *average*, *levering*, *leverage* y *covering*. Finalmente, la schwa es obligatoria en *coverage* y sobre todo en la menos frecuente *Avery*.¹⁵ (El contraste entre *coverage* y *covering* es más que notable, por cierto).

Este gradiente de silabicidad según la frecuencia de uso nos recuerda a la **curva sigmoide** (*S-curve*) que marca la pauta del cambio diacrónico (ver Kiparsky (1995), Aitchison (2001: 89-97)). Gráficamente es:

(9)



La curva indica cómo al principio son pocos los elementos que sufren el cambio, después viene una incorporación masiva de nuevos elementos, y finalmente queda un residuo que tarda más en

incorporarse o ni siquiera se incorpora al cambio. Las condiciones fonotácticas son esenciales pero no lo explican todo, pues.

Por último, el quinto argumento es la existencia de **pares mínimos** con y sin CS. Por ejemplo, el diccionario *Everyman's* da *umble* como /^hʌmbəl/, pero *umbel* es /^hʌmbəl/ [o /^hʌmbel/]. Asimismo, para *javelin* (lanza, jabalina) se prefiere /^hdʒævln/ sobre /^hdʒævəlɪn/, pero con *Javelin* (modelo de coche y avión) la preferencia se invierte. De modo que hay palabras que no tienen CS a pesar de cumplir todos los requisitos para ello, pero sí que la tienen otras palabras muy similares. Es como si la alternancia fuese **arbitraria**. También en checo Scheer (2009: 420) cita *Petr/Petra* ‘Pedro’ (NOMsg., GENsg.) frente a *pater/patro* ‘suelo’ (GENpl., NOMsg.). Sus ejemplos están tomados de Ziková, que atribuye estas diferencias de apariencia arbitraria a una mera variación léxica (la raíz de la variación está en el lexicón).¹⁶ Quizá sea cierto para el checo, desde luego, pero en inglés no podemos conformarnos con la explicación léxica, puesto que el mecanismo es demasiado amplio y productivo.

Y cómo atribuir a una simple variación léxica los **pares idiolécticos**, donde una misma entrada léxica tiene realizaciones distintas de un hablante a otro. O más aún, cuando varía de una preferencia a otra del mismo hablante (por lo que el idiolecto es el mismo: la misma gramática mental, la misma representación). Muy concretamente, en Toft (2002: 141) se observa una variación estadísticamente significativa entre dos preferencias del mismo hablante para la misma palabra. En su diagrama (40), *roughen* tiene una proporción

estadística de silabicidad (número de preferencias con /ŋ/) del 31%, pero es del 47% en su siguiente diagrama, el (41).

A la vista de estos ejemplos, no podemos quedarnos en motivos puramente estructurales para explicar la formación de consonantes silábicas. Las relaciones fonológicas entre segmentos (que son una de las claves de la FR) son necesarias pero no suficientes para dar cuenta de lo que ocurre con las CS.

La alternancia entre schwa y CS

El segundo argumento de la sección anterior decía que en inglés las CS alternan con una forma con schwa. Ahora vamos a profundizar en este aspecto, y para facilitar la tarea empezamos con una distinción terminológica. Cuando sólo es posible la forma con schwa, pero no la CS, hablamos de una **‘forma no-silábica’** o /-əR/ (por ejemplo, *melon* es sólo /'melən/; y *emblem*, *plankton*, *ribbon*, *elegance*, *towel* son también no-silábicas). Cuando sólo es posible la CS pero no la schwa, hablamos de una **‘forma fuertemente silábica’** o /-R/ (*people* y *possible* sólo admiten /ɪ/, *rhythm* es muy preferentemente /'rɪðm/, y *mountain* o *important* tienen una forma fuertemente silábica cuando la /t/ se realiza como [ʔ]). Y cuando sea posible encontrar las dos formas, tanto schwa como CS, hablamos de una **‘forma débilmente silábica’** o /-(ə)R/ (*embosom*, *beacon*, *parent*, *herald*, *minstrel*).

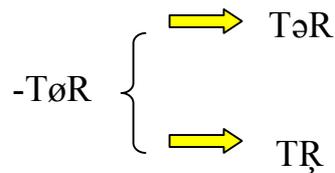
La mayoría de las CS que encontramos en inglés corresponden a la variante ‘débilmente silábica’, donde se alterna entre la schwa y la CS.

La pregunta obligada es cómo se relacionan esas dos variantes, schwa y CS. A lo largo de los párrafos que siguen iremos comprobando la importancia de esta cuestión para el debate entre ‘optimismo’ y ‘pesimismo’, las claves que aporta en cuanto al porqué de las CS.

Para nosotros hay pocas dudas de que la schwa es la **forma subyacente** y la CS es una **forma derivada**. Es decir, que $-\text{əR} \rightarrow -\text{R}$. Con ello se implica que la schwa se está **reduciendo** a CS, puesto que se pierde algo de material fonológico. Este es el planteamiento propio de la hipótesis optimista, que engloba esta reducción dentro de un mecanismo métrico. Además, la idea es coherente con el ‘optimismo’, en tanto que si la schwa es la forma básica, la estructura está bien formada y no hay nada que reparar.

La FR, por el contrario, asume que la forma subyacente es un núcleo vacío, que debe licenciarse o repararse. Pero Scheer (2009: 1) también reconoce, aunque sea implícitamente, que hay una alternancia schwa/CS en alemán cuando habla de “*synchronic (free) variation as e.g. in German leben [leebən] – [leebm̩] ‘to live’ (...)*”. Surge aquí nuestra extrañeza, porque si partimos de un núcleo vacío, la schwa y la CS serían ambas formas derivadas (del propio núcleo vacío), y habría que convenir que el procesamiento se está haciendo en paralelo¹⁷:

(10)



Estrategia reparadora en Scheer (2009):
núcleo vacío subyacente, procesamiento en paralelo.

En la reducción fonológica que nosotros postulamos, en cambio, lo que hay es una secuencia de reglas ordenadas y el procesamiento se realiza en serie:

(11)



Hipótesis optimista: schwa subyacente, procesamiento en serie.

Según hemos anotado en el diagrama, el proceso de reducción puede continuar más allá de la CS, hasta el extremo de vaciar el núcleo por completo y perder esa posición. Esta forma extrema de reducción, que denominamos ‘**supresión**’,¹⁸ se suma a la schwa y la CS para elevar a tres el número de formas que pueden alternar.

Y aún se aprecia una cuarta forma cuando la CS no alterna con schwa, sino con **vocales completas**. Es lo que ocurre en *MAsOn-maSONic*, *MYStery-mysTERious*, *HUNgary-HunGarian*, *MAMmal-mamMAlia*, *NORmal-norMAlity*, o *PREsent-preSENT*, donde la sílaba de la CS aparece como tónica en la otra variante de la oscilación. Aquí nos parece que sería muy forzado partir de una forma subyacente vacía, porque al recibir el acento debería incorporarse una vocal plena, cuyos rasgos son impredecibles. Más difícil todavía: ¿podrían tomarse como básicas las formas [ˈsʌmfɪ] y sobre todo [ˈsʌm·ʔm] y postular que a

partir de ellas o en paralelo a ellas se forma /'sʌmθɪŋ/ (*something*)?

Damos por hecho que esta última es la forma subyacente y las otras se obtienen por reducción¹⁹.

Con las cuatro formas posibles, que son la vocal plena (V), la vocal reducida (v) (que suele ser una schwa), la CS (R̥) y la supresión o cero (∅) obtenemos el proceso de reducción fonológica al completo, con todos los pasos posibles:

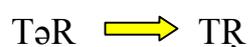
(12)



Por consiguiente, los datos de las alternancias favorecen a nuestra hipótesis frente a la de la FR. Concluimos que en la oscilación entre schwa y CS, la primera debe figurar como forma subyacente y la segunda debe derivarse de ella por reducción. Un dato adicional a favor de la reducción frente a la inserción es que también desde el punto de vista de los mecanismos cognitivos es más sencillo quitar algún dato que rellenar lo que falta.

Una vez que asumimos la reducción y el procesamiento en serie, tratamos de dar cuenta de la variación que nos encontramos al analizar los datos. Podemos enunciar que en las formas fuertemente silábicas el paso de schwa a CS es obligatorio:

(13)



mientras que en las no-silábicas está vedado:

(14)

TəR  TR

Tanto una como otra son del tipo **todo o nada** y sólo admiten un **educto** (*output*). Por el contrario, en las formas débilmente silábicas asumimos que la aplicación de la regla es **opcional**:

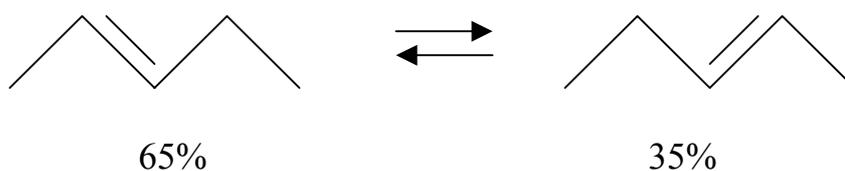
(15)

TəR ( TR)

Y por eso los dos eductos (*outputs*) son posibles. La existencia de más de un educto indica que la forma puede salir de la derivación en uno u otro punto del camino, más o menos indistintamente. El resultado es que dentro de una misma lengua o idiolecto coexisten dos estructuras bien formadas.²⁰

Pero a pesar de estar igual de bien formadas las dos estructuras, en los casos de silabicidad débil no se reparte la preferencia al 50%.²¹ En la práctica tiende a haber preferencia por una de las dos formas (aunque la otra sigue siendo perfectamente posible, claro). Hay entonces una oscilación entre las dos estructuras que tiene porcentajes variables, y que remeda la **resonancia** de los dobles enlaces en la química orgánica:

(16)



Los porcentajes tienen su fundamento: si una estructura tiene mejor porcentaje que la otra, es porque es mejor candidata²². Pero no triunfa sólo la mejor candidata, y esto es esencial; las dos estructuras están igualmente bien formadas desde el punto de vista de las propias reglas. La razón de que una candidata sea ‘mejor’ que otra habrá que buscarla fuera de las reglas estructurales. Cuando tratemos en mayor profundidad las CS/ __V (véase el ‘contexto derecho’ de las CS), veremos la influencia de factores cognitivos (de ‘reconocimiento de formas’) en la preferencia por una u otra forma.

Los porcentajes no sólo varían de una palabra a otra, sino también de un hablante a otro, e incluso de una preferencia a otra. Así pues, hay factores externos a la derivación que influyen en el resultado final, en cuál es la forma que se articula. Un porcentaje más alto puede tomarse como indicio de que la forma en cuestión se adecua más a los módulos con los que interactúa el sistema de reglas fonológicas. En concreto, puede responder a condiciones de la **actuación** (*performance*); por ejemplo, en una situación de cansancio o fatiga, de relajación, de ambiente informal o distendido, se pone en marcha o se da más fuerza a determinado factor que en otra situación.

Aún hay más. En la práctica nos encontramos con todo un espectro de realizaciones, espectro que incluye toda una gama de formas

intermedias entre los extremos de la schwa más resonante y la CS más pura. Casi siempre hay algo de schwa en las preferencias, en distintos grados que podemos asimilar a distintos niveles de reducción. La existencia de toda una gama de realizaciones no nos habla del orden de la derivación, sino de la interacción entre fonología y fonética, y vuelve sobre la idea de que hay más factores que intervienen en la formación de CS, aparte de los puramente estructurales. Este aspecto encaja perfectamente con nuestro planteamiento, y entendemos que lo refuerza. Y al mismo tiempo, no vemos la manera de compatibilizarlo con la hipótesis planteada por la FR.²³

Terminamos así el capítulo dedicado al porqué de las consonantes silábicas. Hemos defendido una hipótesis optimista y holística que atribuye las CS del inglés a una mayor facilidad de procesamiento cognitivo. Las CS inglesas se encuadran en el mecanismo métrico, y los argumentos que hemos aducido ponen en cuestión los motivos de la hipótesis pesimista o reparadora. Finalmente, la alternancia entre schwa y CS es un dato a tener en cuenta en el estudio que estamos abordando.

¹ A esta misma estrategia atribuimos también la emisión de chasquidos (*clicks*) en las lenguas khoisan y en damin (un dialecto iniciático del lardil, véase Blevins 2004:194-97), así como la manera de hablar entre dientes del altiplano andino. Y se nos antoja que las ejectives puedan obedecer a esta misma lógica, al menos en origen.

² En checo se hacen trabalenguas con las CS: *strč prst skrz krk* (literalmente, ‘meter dedo dentro cuello’). Otro trabalenguas es *Smrž pln skvrn zvlhl z mlh* (‘una morilla llena de manchas se humedeció por la niebla’). Además de indicar el carácter marcado que los propios hablantes atribuyen a las CS, los trabalenguas pueden funcionar como **shibboleth**, con matiz nacionalista. Su pronunciación es difícil para un nativo y aún peor para el no-nativo.

Dedicaremos a las CS del checo el último capítulo de esta tesis. Allí pueden encontrarse más detalles sobre estas CS y una comparación con las CS del inglés.

³ La idea de primer y segundo plano, o *foreground* y *background*, está en Langacker (1987).

⁴ Véase Spencer (1996: 283ss.), Cruttenden (2001: 250-51).

⁵ En Fonología de la Rección se habla de **núcleos vacíos**, que necesitan licenciarse. Cuando están a final de palabra o morfema, se licencian paraméricamente. Pero cuando están en interior de palabra, necesitan de otro núcleo que los licencie (mediante rección entre constituyentes, que va de derecha a izquierda). El núcleo licenciador tiene que estar lleno, y situado a la derecha inmediata del N vacío, dentro de la secuencia de núcleos:

$$N_1 \quad \underline{N}_2 \leftarrow N_3$$

A este mecanismo se le llama **Rección Propia** (ver Harris 1994: 191-93), y además de permitir la existencia de núcleos vacíos, impide su proliferación indiscriminada. Como el N licenciador tiene que estar lleno, es imposible tener dos N vacíos contiguos o seguidos. Véase el Apéndice I para más información.

⁶ Los núcleos vacíos sin licenciar deben realizarse fonéticamente: según Harris (1994:192), “*N₂ must therefore be phonetically expressed, either as the latently present @ (yielding fɪdəl) or as a syllabic lateral (fɪdl, through spreading from the following onset)*”.

⁷ Scheer, al igual que Szigetvári, pertenece a un desarrollo reciente de la FR que se conoce como **Fonología CV Estricta**. Scheer defiende la versión **CVCV**; Szigetvári y Blaho, la **VCVC**.

⁸ Las CS/ __V pueden verse a principio de palabra (como en *p'lice* o *s'pport*), en otras formas monomorfémicas (*catalog*, *platinum* o *maintenance*) y también en formas donde la V proviene de flexión o derivación (*softening*, *softener*). El capítulo dedicado al contexto derecho entrará en mayor detalle sobre las CS/ __V, por lo que no vamos a extendernos más por ahora.

⁹ No hay alternancia con schwa en las formas con el sufijo #-ble# o en palabras como *people*, donde la CS queda como la única forma (en dialectos más o menos estándar, se entiende).

¹⁰ En las palabras que ortográficamente acaban en <-gn>, o <-gm>, se decidió elidir la consonante previa a la nasal final, y en las terminadas en <-mn> fue la propia nasal final la que fue suprimida para resolver el choque sonoro. En los derivados morfológicos de estas palabras, sin embargo, la C nasal va seguida por una vocal y entonces deja de haber conflicto sonoro, y por ello las consonantes no se suprimen: compárese *autumn* con *autumnal*, *reign* con *regnal* o *phlegm* con *phlegmatic*. Daremos más detalles en el capítulo 11.

¹¹ Las dos palabras *forgotten* y *forgettin'* aparecen pronunciadas con CS en ‘Your Song’, de Elton John. En cuanto a *fadin'*, Bruce Springsteen la pronuncia [ˈfeɪdn̩] en ‘Streets of Philadelphia’. Nótese que estos gerundios tienen la pronunciación no-estándar con [-n] y no con [-ŋ], de acuerdo con la explicación del capítulo anterior.

¹² En checo existe un sufijo para formar pretéritos, que en uno de sus alomorfos consta solamente de una /j/. Por ejemplo, dentro de la clase I *nést* ‘llevar’ da *nesl*, *péct* ‘cocer’, ‘asar’ da *pekl* y *číst* ‘leer’ da *četl*. En la clase II, *tisknout* ‘imprimir’ da *tiskl*, y entre los verbos irregulares, *říct* ‘decir’ da *řekl*.

Como detalle curioso, la adición del sufijo puede hacer que una palabra tenga dos CS, y nos encontramos con palabras que no tienen ni siquiera una sola vocal: *scvrkl* ‘encogido’, *zmrzl* ‘se congeló hasta solidificarse’, *ztvrdl* ‘endureció’ o ‘endurecido’.

¹³ Y más aún. Hay expresiones que sólo admiten la forma contraída: la expresión coloquial *up an at'em* no puede ser (pronunciarse) *up and at them*.

¹⁴ Hayes (1984: 36) da un ejemplo que encaja en este estudio: *CorNELL HOCkey* → *CORnell HOCkey*, en tanto que *CorNELL athLETics* no sufre cambios, al no haber choque de acentos.

¹⁵ Los datos son del *Everyman's*, por eso utilizamos la ortografía británica.

¹⁶ Scheer (2009: 10-11): *Unlike in German and English where both solutions for final TR# clusters are in free variation, the choice is lexically hardwired in Czech. Given this situation, Ziková considers the Petr – Petra pattern. She argues against a perspective where the syllabicity of consonants is understood as a repair strategy for orphan empty nuclei. (...) The solution favored by Ziková therefore anchors the syllabicity of the final R in the lexicon.*

¹⁷ Evidentemente, esto no es nada malo en sí, pero tiene consecuencias teóricas, y hay que tenerlo en cuenta

¹⁸ La supresión sucede en contexto vocálico /__V, según veremos al hablar del contexto derecho.

¹⁹ En este ejemplo interviene la frecuencia de uso y la posibilidad (fácil, accesible) de reconocer *something* como la palabra original, previa a la reducción. Todos los oyentes reconocen *something* en la pronunciación [ˈsʌm.ɪŋ] de Bruce Springsteen en la canción ‘Sad Eyes’.

²⁰ Esta forma dinámica de plantear las estructuras puede ser una innovación respecto a la teoría fonológica al uso, que se ha entendido de forma más estática.

Al margen de lo expuesto, no vemos cómo explicar la existencia de formas no-silábicas, fuertemente silábicas y débilmente silábicas únicamente a partir de motivos puramente estructurales, como se hace en la FR.

²¹ En las formas ‘todo o nada’ los porcentajes de aparición de /-əR/ y /-R/ son de 100/0 (o viceversa).

²² Los porcentajes sirven para cuantificar el grado de preferencia, aunque ya hemos visto que los propios porcentajes oscilan de una preferencia a otra.

La distribución de porcentajes en la oscilación es un buen complemento de la curva sigmoide de Kiparsky. Por ejemplo, en *everything* o *everybody* el porcentaje de ‘cero’ está cerca del máximo, y por tanto el de ‘schwa’ está cerca del 0%. En *every* el porcentaje de ‘cero’ es algo menor, pero aún es bastante alto. En *several*, los porcentajes están más repartidos entre ‘schwa’ y ‘cero’, y en *levering* o *leverage* probablemente la schwa tenga el porcentaje más alto, y en *Avery* el porcentaje de ‘schwa’ estará cerca del 100% cuando no se pronuncie en el ámbito local, esto es, cuando no sea una palabra habitual para esos hablantes.

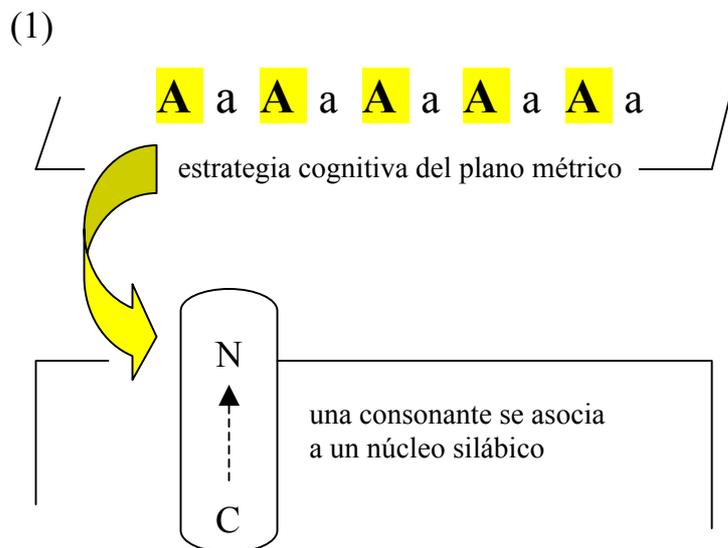
²³ También nos planteamos cómo tratar un espectro continuo de realizaciones desde el punto de vista de la Teoría de la Optimidad. Igual que nos parece muy interesante para este marco teórico la posibilidad de que puedan triunfar dos candidatos, dado que en esta teoría hay una jerarquía de restricciones que lleva al triunfo de un solo candidato.

Las **restricciones** a las que se alude en la Optimidad bien pudieran ser las fuerzas externas que interactúan con la representación fonológica, incluyendo, como hemos señalado, los factores de la actuación. En conjunto son varios factores los que desplazan el equilibrio en uno u otro sentido. La innovación que introducimos ahora supone que en un momento puede prevalecer cierto factor, pero en otro momento puede ser otro factor el que gane, de manera que la **jerarquía de restricciones** no es del todo fija.

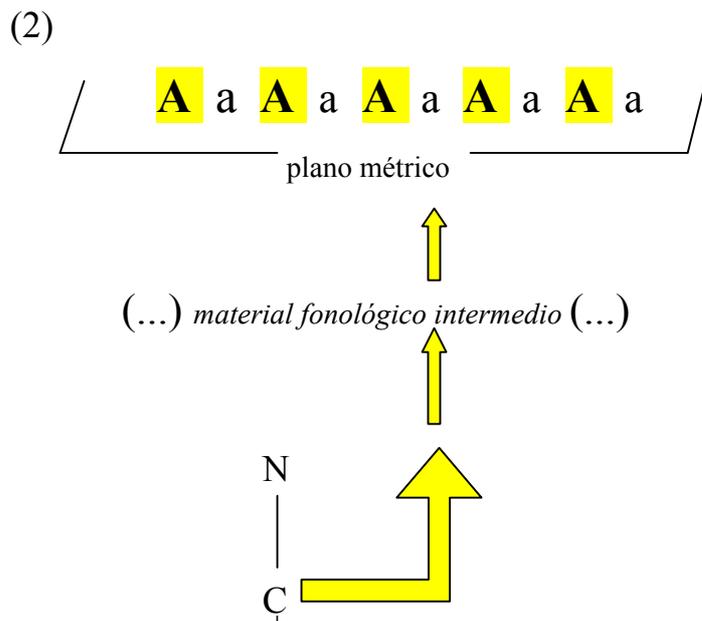


3 Cómo se forman las CS (I)

Si en el capítulo anterior vimos *por qué* se forman las CS del inglés, ahora vamos a dedicar varios capítulos a entender *cómo* se forman esas CS. Una vez que tenemos el motivo para formar las CS, el paso siguiente es implementar esa formación: cuáles son los mecanismos estructurales que la hacen posible. Nuestro modelo sostiene que las CS inglesas responden a una estrategia cognitiva localizada en el plano métrico, y ese es un lugar muy alejado del plano segmental, donde se articulan las CS:



Si la estrategia opera sobre un lugar tan lejano, se debe sin duda a que la formación de una CS tiene a su vez consecuencias sobre la estructura métrica:



La formación de una CS afecta a la estructura métrica

Como el inductor es el material métrico, damos por hecho que toda la estructura tiene que estar construida antes de poner en marcha la formación de una CS. Este requerimiento es aún más fuerte si la estructura se construye de abajo arriba, como hemos asumido en la introducción. Y sobre todo, es necesario que el cambio que una CS acarrea pueda transmitirse hasta el lugar más alto de la estructura, como damos a entender en el diagrama. Desde esta perspectiva, lo que estamos proponiendo es que la formación de una CS tiene efectos sobre toda la arquitectura fonológica¹.

Por eso nuestro enfoque va a ser holístico. Con esta directriz vamos a empezar por revisar la manera en que se organiza el espacio fonológico. Nuestro modelo revisado busca una integración de todos los niveles de la arquitectura fonológica, desde el más alto hasta el más bajo.

Cómo organizamos el espacio

Para el estudio hemos elegido un modelo derivacional de reglas ordenadas². Asumimos que la estructura se construye de abajo arriba, en formas cada vez más complejas y sin volver atrás. Es más, la información de los estratos inferiores va ascendiendo a los estratos más altos a medida que la derivación progresa, y la información que hay abajo siempre debe estar disponible en los niveles más altos. Se mantiene así la integración que venimos defendiendo.

Todos los estudios que se han hecho en fonología hasta la fecha, que nosotros sepamos, centran su atención en una parte de la derivación, y dejan de lado el resto. Por ejemplo, unos se centran en los segmentos, otros en los rasgos subsegmentales, otros en la estructura silábica, y otros en la métrica. Pero el enfoque que proponemos tiene en cuenta todos los niveles de la estructura, y nuestro estudio se irá deteniendo en cada sector a su debido tiempo, para ver los detalles de cómo se forman las CS en cada nivel.

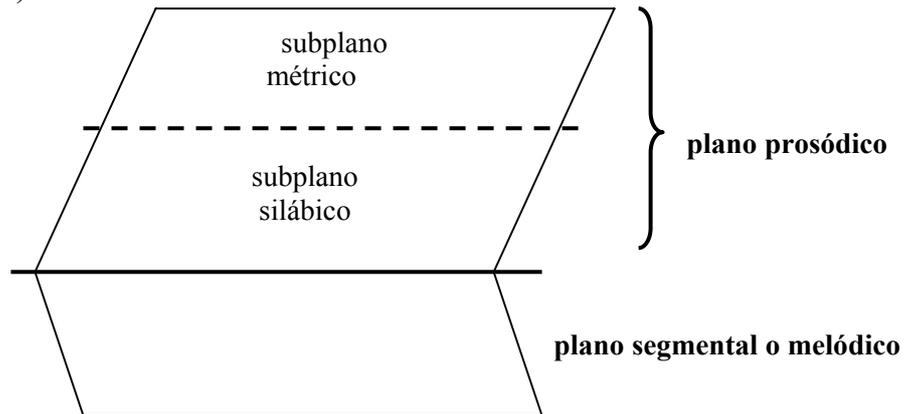
Queremos empezar por dar un marco general sobre el que se edifica toda la arquitectura fonológica. El **espacio** que pasamos a definir se organiza en **planos**, según el patrón de la **Fonología Autosegmental**, que en general se ha incorporado a la fonología no-lineal posterior al SPE (*Sound Pattern of English*). En esta distribución, los segmentos y sus rasgos o elementos constituyentes se localizan en el plano segmental, los nodos silábicos están en el plano silábico, y la estructura métrica está en el plano métrico. (Insistimos en que los

estudios suelen concentrarse en uno solo de los planos, o en dos de ellos a lo sumo.)

Los planos segmental y silábico se articulan a través de una **gradilla de tiempo o esqueleto autosegmental**. Pero en lo respectivo a los planos más altos se aprecia una desconexión entre los nodos silábicos y los constituyentes métricos. Cuando los estudios se han centrado en la estructura silábica, el nodo silábico servía como “techo” del plano silábico. Y ese mismo nodo silábico era el que formaba el “suelo” del plano métrico (la ‘línea 0’) en los estudios de Fonología Métrica. Pero como cada estudio se concentraba en un sector de la estructura, siempre ha dado la impresión de que “techo” y “suelo” no se tocaban. Por eso ahora queremos dar una representación unificada, con la cual definimos un ámbito general para los constituyentes fonológicos.

Definimos entonces un espacio fonológico constituido en total por **dos planos autosegmentales**, que vamos a llamar ‘**melódico**’ o inferior y ‘**prosódico**’ o superior. En realidad son los mismos planos que se vienen manejando en Fonología post-SPE, y lo que vamos a cambiar es la manera de integrarlos y articularlos. El plano melódico contiene los segmentos fonológicos³, como siempre, pero los nodos de sílaba y las unidades métricas quedan integradas en un solo plano, que es precisamente el plano prosódico. Proponemos que este plano superior esté articulado en **dos subplanos**: en el ‘**subplano silábico**’ están los nodos de la estructura silábica (inicio, núcleo, coda, rima y sílaba), y por encima se sitúa el ‘**subplano métrico**’, con los pies y las unidades superiores de la jerarquía prosódica:

(3)



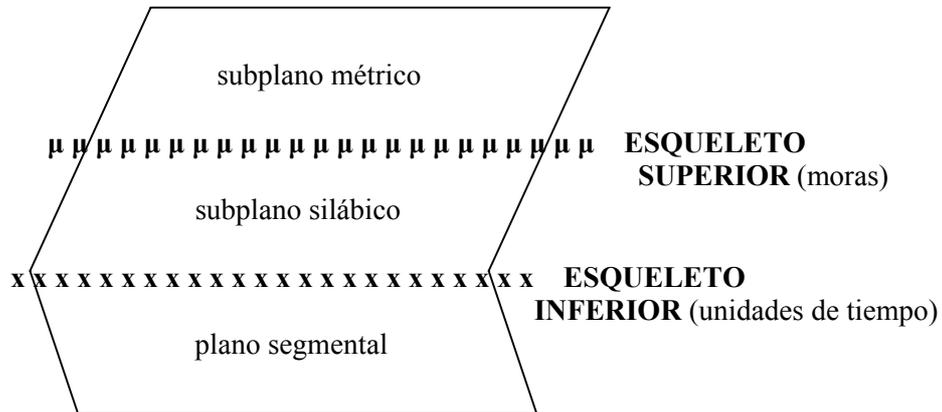
No es muy distinto de lo que había, salvo por el hecho de que el espacio ha quedado reducido a sólo dos planos, y así nodos silábicos y estructura de pies quedan todos incluidos en un mismo plano. Así es cómo deja de haber desconexión entre ‘techo’ de la estructura silábica y ‘suelo’ de la estructura métrica⁴.

Otra novedad que introducimos es la manera de articular los ámbitos. El plano melódico se articula con el plano superior mediante un **esqueleto de unidades de tiempo**, como siempre. Pero ahora el esquema se traslada también a la articulación entre los dos subplanos del plano prosódico. Eso sí, dejamos de tomar los nodos de sílaba como base de las unidades métricas (en la formación de pies); los nodos silábicos van a quedar incluidos DENTRO del subplano correspondiente, y la articulación entre lo silábico y lo métrico se hace mediante otro esqueleto autosegmental, al que vamos a llamar **‘esqueleto superior’**, por oposición al **‘esqueleto inferior’**.

La particularidad del esqueleto superior es que no está formado por unidades de tiempo exactamente, sino por unidades de peso, es decir,

por '**moras**'. En una palabra, el esqueleto superior no es ni más ni menos que la **gradilla moraica**:

(4)

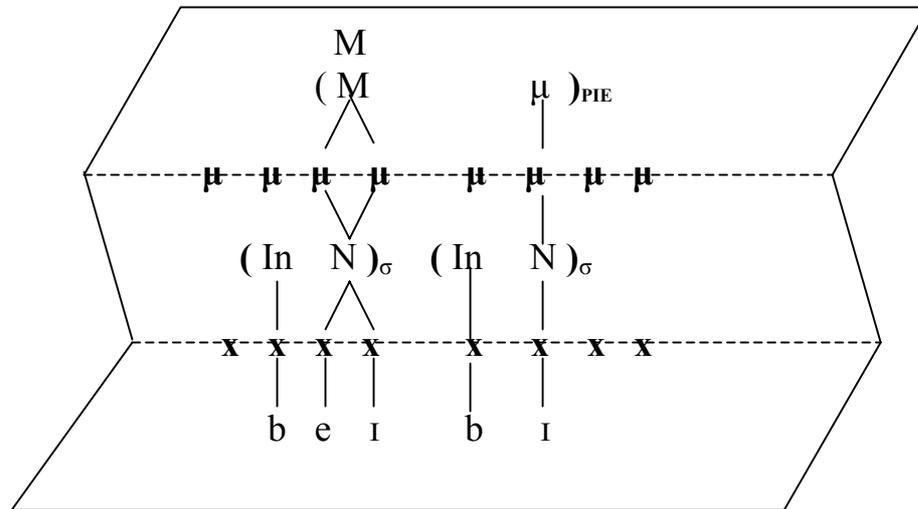


Hemos planteado así el nuevo modelo porque entendemos que el peso no incumbe a los segmentos, sino a las unidades métricas. En concreto, el **peso por posición** (Hayes 1989: 356-7) tiene en cuenta los nodos de la rima (núcleos y codas) al tiempo que descarta los inicios. Nosotros pretendemos totalizar los dos componentes del **peso métrico**, que son la duración⁵ (consignada en el esqueleto inferior) y la posición⁶ (consignada en los rótulos silábicos); y a tal efecto vamos a situar las moras por encima de los rótulos silábicos. Creemos que así puede zanjarse el eterno conflicto entre gradilla de tiempo y gradilla moraica, simplemente ubicando cada una de ellas en un sitio diferente, aunque conservando las analogías necesarias.

Será más fácil entender cómo funciona el nuevo modelo si lo plasmamos en un ejemplo concreto, y de paso nos encontraremos con dificultades nuevas, que darán más profundidad al estudio. Vamos a hacer la transcripción completa de una palabra del inglés, con todas las

proyecciones de los planos. Elegimos una palabra común: *baby*, que es /^hbeɪbɪ/:

(5)



En el diagrama tenemos los segmentos colocados en el plano melódico: b-e-ɪ-b-ɪ. Esos segmentos se asocian a unidades de tiempo del esqueleto inferior (las aspas 'x'), y con esa articulación se asocian a su vez a los nodos silábicos correspondientes, que están en el subplano silábico: inicios (In), núcleos (N) y codas (Co). A su vez, los nodos N y Co se asocian a moras (μ) del esqueleto superior (no así los inicios), y desde ahí proyectan (lo hemos representado como M y μ) al subplano métrico, que es donde se forman los pies. Queremos aclarar que en los esqueletos hemos colocado unidades de más (unidades de tiempo y moras), para dar idea de continuidad⁷ y mostrar cómo hay constituyentes que no se asocian al esqueleto.

Nótese que en el subplano silábico sólo hemos utilizado los nodos básicos (Inicio, Núcleo y Coda), pero no sus proyecciones (Rima y Sílabas). En este punto no podemos dar una opinión definitiva sobre

cómo proyectar los nodos. Puede utilizarse la representación habitual, donde se proyecta una estructura arbórea en la que se forman los nodos más altos (Rima y Sílabas) dentro del propio plano silábico; o bien la que hemos usado en el diagrama, agrupando los nodos mediante **paréntesis rotulados**, con lo que damos cabida a que las proyecciones silábicas pudieran realizarse en un plano perpendicular. Una tercera alternativa sería prescindir de rótulos y proyectar mediante **estructura escueta** (*Bare Phrase Structure*)⁸, y podrían usarse subíndices. Hemos aplicado un criterio análogo en el subplano métrico: los pies aparecen como agrupaciones con paréntesis rotulados, aunque igualmente podrían representarse como proyecciones sintagmáticas⁹.

Una ventaja de la nueva representación es que el peso está reflejado en el esqueleto superior. Tradicionalmente, la línea básica para la métrica era una secuencia de sílabas (nodos σ), sin que se hiciera constar en ella cuál era pesada y cuál era ligera. Pero al articularse los planos a través del esqueleto superior, la representación y traslación del peso deja de ser arbitraria o estipulativa. La gradilla moráica recoge el peso (en forma de asociaciones) directamente de los nodos silábicos, y lo traslada (en forma de nuevas asociaciones) al componente métrico. Es exactamente la misma función que ha venido haciendo el esqueleto inferior con la duración de los segmentos.

El diagrama anterior no ha pretendido recoger todas las posibilidades del peso, porque habría sido una tarea excesiva. Hacen falta varios diagramas más detallados para cubrir las distintas especificaciones, como veremos a continuación. En inglés, que es la lengua que estamos estudiando, se hace una distinción entre sílabas ligeras y pesadas. Las

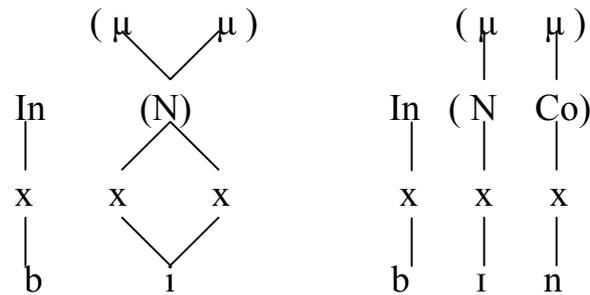
sílabas **ligeras** tienen solamente una mora o unidad de peso, porque sólo tienen un núcleo corto; ese núcleo está asociado a una sola unidad de tiempo del esqueleto inferior, y eso es lo que a su vez se proyecta al esqueleto superior:

(6)



Casi está de más decir que las sílabas ligeras no tienen coda. Por el contrario, las sílabas **pesadas** tienen dos moras, que pueden provenir de un núcleo pesado (con dos unidades de tiempo) y/o de una coda¹⁰. El núcleo pesado, como es sabido, puede formarlo una vocal larga o un diptongo (por ejemplo, *bee* /bi:/ o *by* /baɪ/). Como ejemplos de núcleo ligero más coda tenemos *bin* /bɪn/ y *mint* /mɪnt/, y como ejemplos de núcleo pesado más coda, *mine* /maɪn/ y *mind* /maɪnd/. Las sílabas pesadas tienen una rima pesada, en definitiva; sus nodos N y/o Co están asociados a dos o más unidades de tiempo del esqueleto inferior, y así se proyecta al esqueleto superior:

(7)

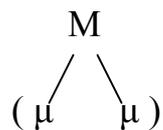


Silabas pesadas: núcleo pesado (*bee*) y coda (*bin*)

N.B.: Los paréntesis no indican opcionalidad, sino agrupación de constituyentes.

Una vez en el esqueleto superior, las moras también se agrupan en paréntesis. El diagrama muestra cómo se conservan los paréntesis o agrupaciones que vienen ya dados desde la rima en el subplano silábico; en realidad, lo único que hacen es proyectarse, sin necesidad de estipular nada. Cuando dos unidades de la gradilla moraica (que hemos representado con la letra μ) están agrupadas, pueden proyectar a su vez a un nivel más alto, y forman una unidad que vamos a denominar ‘**supermora**’, representada por la letra M:

(8)



Adviértase que la gradilla moraica funciona como la **línea 0** de la Fonología Métrica (véase Hayes 1995). La supermora M funciona como núcleo del pie (trocaico, en este ejemplo), y como tal y es capaz de formar un pie por sí sola, puesto que tiene el peso mínimo necesario. Así se forma un ‘**troqueo simple**’: ¹¹

(9)

$$\left[\begin{array}{c} M \\ (\mu \quad \mu) \end{array} \right] \begin{array}{l} \text{PROYECCIÓN ('LÍNEA 1')} \\ \text{GRADILLA MORAICA ('LÍNEA 0')} \end{array}$$

Y por supuesto que la supermora puede también asociarse con otra mora simple para formar un pie:

(10)

$$\left[\begin{array}{c} M \\ M \quad \mu \end{array} \right]_{\text{PIE}} \begin{array}{l} \text{PROYECCIÓN ('LÍNEA 2')} \\ \text{PROYECCIÓN ('LÍNEA 1')} \end{array}$$

$\begin{array}{c} \wedge \\ (\mu \quad \mu) \quad | \\ \mu \end{array}$

GRADILLA MORAICA ('LÍNEA 0')

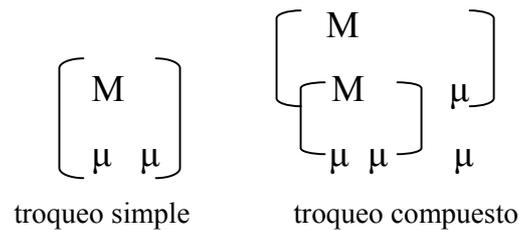
A este pie lo vamos a denominar **'troqueo compuesto'**,¹² que en realidad se basa en el mismo patrón estructural que el troqueo simple. En efecto, los dos tienen el esquema:

(11)

$$\left[\begin{array}{c} X \\ x \quad x \end{array} \right]$$

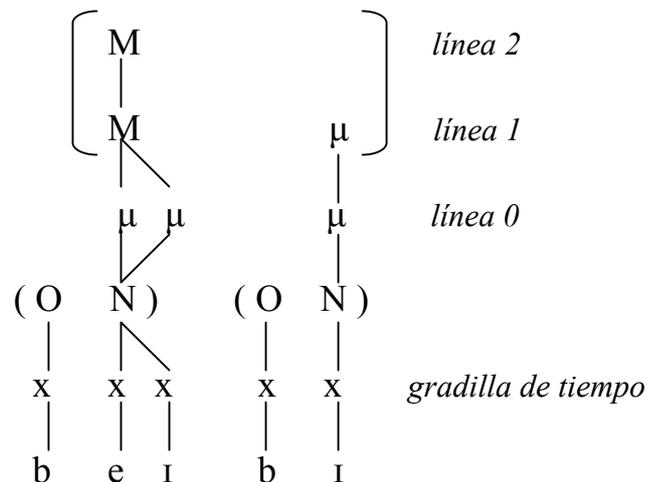
Es decir, que aplicando un solo mecanismo pero en dos lugares distintos, obtenemos las dos estructuras trocaicas. Y más todavía, en el troqueo silábico se itera la proyección, con la cabeza a la izquierda. Vamos a ponerlos juntos para que se aprecie mejor:

(12)



La manera en que está planteado el esqueleto superior tiene tres ventajas. La primera, como ya sabemos, es que la proyección de pies no es arbitraria; hasta ahora se colocaba una secuencia de sílabas (la ‘línea 0’) en la gradilla métrica, pero había que hacer constar de alguna manera cuál era pesada y cuál ligera. Con el esqueleto moraico, el peso simplemente se proyecta, y lo único que hay que hacer es trasladarlo a niveles más altos. La segunda es que el acento recae directamente sobre el núcleo¹³, puesto que todo el aparato es una sucesión de proyecciones. Para el ejemplo de *baby* es¹⁴:

(13)



La tercera ventaja es que se sigue un único mecanismo para los dos esqueletos, superior e inferior. Conforme a la **navaja de Ockham**, es más **económico** no multiplicar las entidades o las reglas, y si basta con

trasladar el mismo modelo a otro sitio para obtener los resultados que corresponden a ese otro nivel, es preferible hacerlo así.

Lo único que estamos utilizando, en definidas cuentas, es un mecanismo de proyección. Los segmentos melódicos se proyectan hacia el subplano silábico a través del esqueleto inferior, y la información relevante llega a niveles más altos de la estructura. Del mismo modo, la información del subplano silábico se proyecta hacia el subplano métrico, utilizando la mediación del esqueleto superior, y no se “pierde” nada por el camino. La disposición del espacio es crucial en el modelo que estamos proponiendo: la distribución de los elementos, y también el lugar en el que operan las reglas fonológicas. En este sentido, el *cómo* es también el *dónde*.

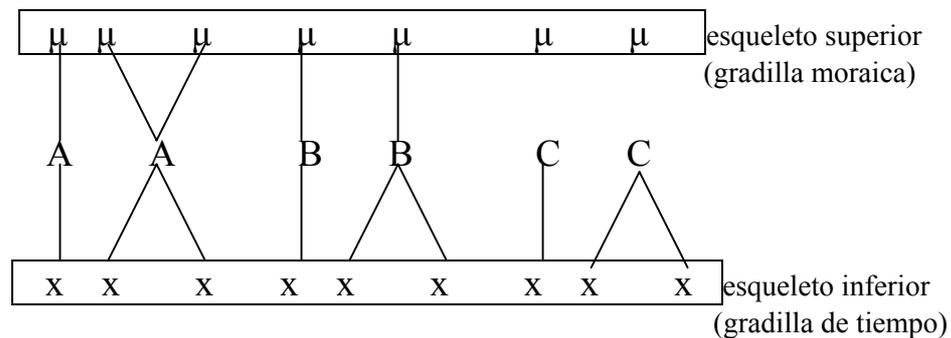
Sensibilidad al peso en función de la transparencia

Hasta este momento nos hemos centrado en el inglés para trazar las líneas maestras del modelo, pero también es necesario dar cuenta de lo que ocurre en otros tipos de lenguas. Así podemos probar la capacidad predictiva del modelo. Existe una distinción básica entre lenguas sensibles y lenguas no-sensibles al peso (Ewen & van der Hulst 2000: 223-25); y dentro de las lenguas sensibles se observan diferencias respecto a lo que se cuenta como peso (2000: 133-36, 150-54).

Pues bien, el modelo autosegmental revisado puede formular todas esas diferencias a través de un mecanismo sencillo al que vamos a llamar ‘**transparencia**’. Todo se reduce a cuáles son los nodos o constituyentes que pueden trasladar su información a niveles más altos

y cuáles no pueden hacerlo. A un constituyente que puede proyectar toda su información lo llamamos ‘**transparente**’, y al que no proyecta nada lo llamamos ‘**opaco**’. Asimismo, llamamos ‘**semitransparente**’ al que proyecta sólo una unidad, con independencia de cuántas reciba. Veámoslo gráficamente:

(14)



A=nodos transparentes; B= nodos semitransparentes; C= nodos opacos
 (lo proyectan todo) (proyectan sólo uno) (no proyectan)

Las diferencias en el grado de transparencia producen diversas combinaciones. Todos los inicios son opacos¹⁵, por definición¹⁶: la cantidad que reciben del esqueleto inferior jamás se proyecta al esqueleto superior. Los núcleos pueden ser transparentes o semitransparentes. Y las codas son coadyuvantes de los núcleos en la aportación de peso; si aportan una unidad de peso (como máximo), son semitransparentes, y si no pueden aportar ninguna, son opacas.

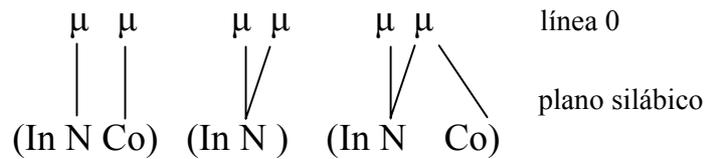
Asumimos que el máximo de peso que una rima puede aportar es de dos unidades, o sea que una rima puede ser pesada a lo sumo¹⁷. Recordemos la distinción entre peso (que está en la gradilla moraica) y cantidad (que está en las unidades de tiempo). Desde el punto de vista de la cantidad sí que podemos hablar de rimas **superpesadas**

(hasta cuatro unidades en *mind*), pero por encima del esqueleto superior nunca vamos a distinguir más allá de lo pesado (dos moras); es una forma de semitransparencia, ya que la proyección al esqueleto superior no se corresponde enteramente con la cantidad recibida del esqueleto inferior.

Veamos algunas diferencias entre lenguas en función de la transparencia. En las **lenguas insensibles al peso** cada sílaba proyecta una única línea de asociación hacia el esqueleto moraico. Esta asociación biunívoca (una sílaba, una mora) es siempre la misma, con independencia del material que haya por debajo del nodo σ . Por eso estas lenguas pueden concebirse como semitransparentes respecto a la sílaba (el nodo σ); o si acaso, como semitransparentes respecto al núcleo (el nodo N) y con las codas opacas.

Por el contrario, en las **lenguas sensibles al peso** la proyección no es siempre biunívoca (sílabas/moras). En concreto, las **lenguas con peso en la rima** (*rhyme-weight languages*, o bien *branching rhyme* en Hayes, 1995: 51-52) pueden proyectar desde los nodos N y Co. Eso sí, respetando siempre el principio de que la proyección principal corresponde al N, donde la Co es sólo coadyuvante, y contando con que sólo aceptamos un máximo de dos moras por sílaba. Este es el caso del inglés¹⁸, donde ya hemos visto en qué consisten las sílabas ligeras y las sílabas pesadas:

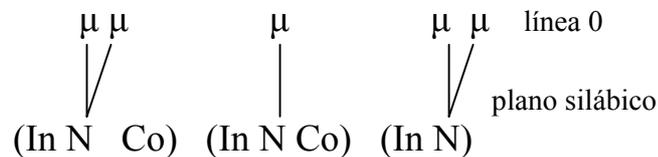
(15)



el inglés es una lengua con peso en la rima

Hay otras lenguas, como el selkup, donde el peso proviene exclusivamente del núcleo [*nucleus-weight languages* según Ewe & van der Hulst (2000: 134), que son las de *branching nucleus* en Hayes (1995: 51-52)], por lo que sólo son pesadas las sílabas con núcleo pesado. En cambio, la presencia de una coda no tiene relevancia en el recuento del peso (es opaca, en los términos que estamos proponiendo):¹⁹

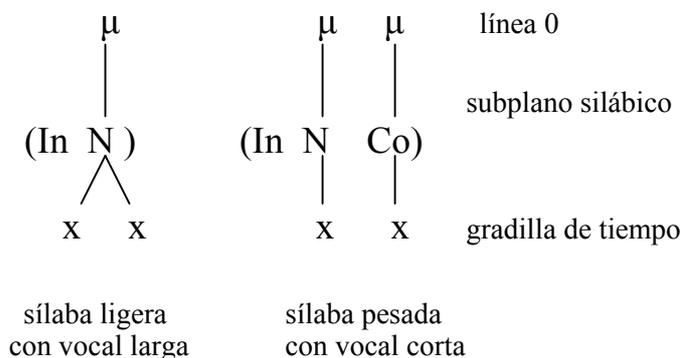
(16)



el selkup es una lengua con peso nuclear

Un tercer tipo de lenguas sensibles al peso está representado por el holandés. En esta lengua, las vocales largas se realizan fonológicamente como tales vocales largas, pero sus dos unidades de tiempo sólo pesan como una cuando se asigna el acento de palabra. Dicho en los nuevos términos, el núcleo es siempre semitransparente. Para que una sílaba en holandés cuente como ‘pesada’ a efectos de acentuación, debe estar cerrada por una coda (Ewen & van der Hulst (2000: 135) lo denominan **coda languages**).²⁰ El nuevo modelo recoge perfectamente todos estos matices:

(17)



en holandés sólo son pesadas las sílabas cerradas

De hecho, podemos considerar que las lenguas tienen ‘**parámetros de transparencia**’ que regulan cuáles son los nodos que proyectan y cómo lo hacen (es decir, como transparentes o como semitransparentes).²¹ Análogamente al clásico parámetro para nodos de subyacencia (que asignaba sintagma complementante SC al inglés pero sintagma flexivo SF al italiano), lo que se parametriza es una categoría en concreto, no es simplemente ‘sí’ o ‘no’.²²

En definitiva, las **lenguas con peso en la rima** son totalmente transparentes, tanto en el nodo N como en el nodo Rima; por eso pueden recibir el peso tanto del núcleo como de la coda. Por eso es la categoría menos restringida. Por su parte, las **lenguas de peso nuclear** sólo tienen el N transparente, pero la coda es opaca. Y en las **lenguas de coda**, la rima es transparente, pero el N es semitransparente (sólo deja pasar una unidad).

En este capítulo hemos abordado una revisión del espacio fonológico, donde se integra todo el material fonológico de manera coherente, con un solo espacio y un solo sistema de articulación y proyecciones.

Gracias a esta unificación resulta más fácil conectar las CS con la estructura métrica que las pone en movimiento, según habíamos planteado al principio del capítulo.

¹ Hemos visto cómo en cambio el modelo de la FR agrupaba en los nodos de la estructura silábica tanto la razón de ser de las CS como su organización estructural. Todo quedaba localizado en un mismo ámbito.

² Hemos optado por la derivación frente al modelo de filtros o restricciones que se aplica en la Teoría de la Optimidad. Por supuesto que el debate entre los dos modelos no está zanjado, y es probable que la fonología y el lenguaje utilicen los dos sistemas (es decir, tanto reglas como restricciones; y procesamiento tanto en serie como en paralelo). Nuestra elección tiene algo de opción personal, pero también tiene en cuenta los buenos resultados que el modelo derivacional ha dado en la investigación que la FR ha seguido en torno a las CS.

Sí queremos comentar, lejos de cualquier descalificación, el dilema que las CS débilmente silábicas parecen plantearle al modelo de restricciones. Una posible objeción está en que la schwa y la CS son dos estructuras bien formadas, en tanto que la Optimidad sólo admite un ganador. Otra objeción viene de que con el procesamiento en paralelo no se aprecia si una de las formas se deriva de la otra, ya que las dos sirven como formas terminales o definitivas. El problema es que si aceptáramos que la CS no proviene de la forma con schwa, estaríamos dando al Generador (y a los filtros restrictivos, de paso) la capacidad de producir una CS insertada directamente en el nodo N, en una suerte de “todo vale” que a nuestro juicio tiene muy poco poder explicativo.

³ Los segmentos están formados a su vez por rasgos o elementos fonológicos.

⁴ Por cierto que esta manera de organizar los constituyentes remeda la división entre **árbol léxico** y **árbol funcional** que se ha hecho en sintaxis. Remitimos a Haegeman (1994), capítulos 2 y 11 para más información.

En términos generales, tanto el árbol léxico de la sintaxis como el plano melódico de la fonología tratan de lo ‘sustancial’. En fonología, lo sustancial es el material fónico, que va articulado en múltiples planos autónomos.

Y lo funcional en sintaxis o el plano prosódico (silábico + métrico) en fonología va destinado a organizar la información, darle estructura y prepararla para su procesamiento y para interactuar con los módulos de interfaz, a los que se remite esa información lingüística.

⁵ En efecto, lo que se codifica en el esqueleto inferior no es el PESO, sino la CANTIDAD o DURACIÓN de los segmentos melódicos.

⁶ Es lo que Hayes (1989) llama **peso por posición**. Por ejemplo, se asume que los inicios no tienen peso, por mucho que tengan cantidad (y más aún cuando el inicio es complejo o ramificado). Por consiguiente, el peso se constituye mediante la combinación de un componente cuantitativo, que es la cantidad segmental que recoge el esqueleto inferior (la *timing tier*), y otro componente cualitativo, que depende de los nodos de la estructura silábica y da lugar al peso por posición. Este es el peso moraico, que es el que se tiene en cuenta para la asignación del acento en las lenguas que son sensibles al peso.

⁷ En geometría analítica, una recta puede definirse como la intersección de dos planos.

⁸ Véase el capítulo de introducción para el concepto de estructura escueta.

⁹ También aquí se mantiene una coherencia holística, porque en Fonología Métrica se prefiere la notación de aspas y paréntesis (tipo Hayes) a la arbórea, mientras que los nodos silábicos se proyectan en forma de árbol.

¹⁰ Vamos a ceñirnos a los términos tradicionales y evitar hablar de inicios finales (seguidos de núcleos vacíos) como se hace en la FR, con la intención de no complicar el modelo en exceso.

¹¹ El troqueo simple es similar al ‘troqueo moraico’ de Hayes 1995, pero no idéntico. Desarrollaremos esta cuestión en el Apéndice III, por lo que no abundamos más en ella por ahora.

¹² El troqueo compuesto remeda el ‘troqueo silábico’ de Hayes 1995, pero hay alguna diferencia. De nuevo remitimos al Apéndice III de esta tesis para un análisis más detallado y explicativo.

Nótese que el troqueo compuesto consta de tres elementos, pero los constituyentes se agrupan de dos en dos. Es decir, que respetamos la tendencia a la binariedad que observa Hayes 1995:

“Much recent theorizing in generative grammar focuses on the idea of locality: we obtain interesting and valid predictions by constraining rules to apply within bounded domains. In phonology, the principle of locality often takes the form of limiting what can be counted: a reasonable conjecture is that phonological rules can count only to two.” (1995: 307).

O sea que vamos a evitar los pies dactílicos y a limitarnos a los pies binarios para el análisis. Cuando haya una secuencia de una parte fuerte seguida de dos débiles, podemos agrupar dos de esas partes en un grupo. Por ejemplo, Hayes analiza en términos binarios la métrica del cayuvava, y nosotros también pero preferimos evitar el recurso a la extrametricidad de Hayes, y en su lugar asignamos agrupaciones de moras. Este es nuestro análisis para *kihíbere* ‘yo corrí’:

M
(μ μ) μ
ki hi be re

No hay ningún elemento extramétrico, sino que se ha formado un troqueo. Las dos sílabas agrupadas sirven para formar la supermora, porque al ser sílabas abiertas no tienen peso suficiente para formarlas de una en una. Y la sílaba final forma el ‘resto’ que completa el troqueo. El análisis se puede trasladar igualmente a los otros ejemplos: *ikitaparerepeha* ‘el agua está limpia’ y *čaadirobošururuče* ‘noventa y nueve’, primer dígito; el segundo dígito de ‘noventa y nueve’ es *čaadairo boi*.

¹³ Szigetvári & Scheer (2005) afirma que el acento es una propiedad nuclear, puramente vocálica.

¹⁴ Nótese que en el plano melódico hemos escrito /beɪbɪ/, sin la tilde que marca el acento principal de palabra (las transcripciones al uso siempre la incluían). La única razón para ponerla en un nivel tan bajo sería para hacer constar que ya viene especificada desde el léxico.

¹⁵ La verdad es que hay que plantearse el peso de los inicios si queremos mantener la afirmación de que en inglés todas las sílabas tónicas son al menos dimoraicas (por eso Hammond (1999: 202) distingue entre *beck-on* y *ba-con*). En el capítulo 9 daremos una posible explicación de este fenómeno, de cómo funciona, cuando hablemos de la ‘captura de mora’.

¹⁶ “Por definición” es como se viene asumiendo. Dice Szigetvári & Scheer (2005) que los inicios no tienen peso porque nunca se ha observado tal cosa. Y que a partir de esa observación se llega a una definición circular de qué elementos tienen peso y cuáles no.

Con nuestro modelo podemos ser más explicativos, pensamos. La falta de peso del inicio puede atribuirse a la opacidad (o falta de transparencia), pero si esa condición no se fundamenta, seguirá siendo pura estipulación.

Pero gracias a la integración de los niveles de la estructura, podemos derivar la falta de peso métrico de las consonantes de inicio a partir de dos factores. Uno es la independencia fonotáctica que tienen respecto al núcleo. Tengamos en cuenta que el peso es una propiedad esencialmente nuclear, y que al fin y al cabo es una proyección (a un nivel más alto) del núcleo o de la rima (teniendo en cuenta que sí hay relación fonotáctica entre núcleo y coda).

En segundo lugar, la falta de peso del inicio puede derivarse de la propia estructura silábica. El inicio está en la posición de **especificador**, y asumimos que desde ahí no puede proyectar al nivel métrico. Sin embargo, el núcleo tiene que proyectar necesariamente, por su misma condición de cabeza de la estructura, y la coda puede hacerlo o no (determinado paramétricamente) en función de que la **proyección intermedia** o el propio **complemento** se tengan en cuenta.

Y desde luego, los dos argumentos están íntimamente conectados. La estructura silábica, con su cabeza, complemento y especificador, así como proyección intermedia (la rima) son un reflejo de las relaciones fonotácticas que se contraen entre los componentes de la sílaba.

¹⁷ En Hayes (1989: 355) vemos varios segmentos asignados a una sola mora (en *ant*), y en Hayes (1995: 53, diagrama 40) vemos que /ta:t/ consta solo de dos moras.

¹⁸ Hay varias muestras de la sensibilidad al peso del inglés:

Uno: el acento tiende a recaer sobre una sílaba pesada.

Dos: las sílabas átonas se reducen, y tienden a formarse pies isocrónicos.

Tres: no se permite la formación de *palabras monomoraicas: los monosílabos son al menos dimoraicos.

¹⁹ Esto es, en este tipo de lenguas una sílaba cerrada sólo es pesada si su núcleo es pesado. Pero si el núcleo es ligero, la sílaba es también ligera, monomoraica.

Hayes (1995: 52) dice que “*languages like Latin allow a postvocalic consonant within the syllable to bear a mora (symbolized /μ/), whereas languages like St. Lawrence Island Yupik do not*”, lo cual podemos traducir como falta de transparencia métrica a las consonantes de la coda.

Por cierto que Hayes (1989: 356) incluye el lardil entre estas lenguas: *We must also account for languages like Lardil, where CVC is light*. Lo que pasa es que Hayes tiene una explicación algo diferente de la nuestra: *I assume that such languages have no Weight by Position rule, so that the final consonant is made a daughter of the final mora*.

Y añade un detalle interesante: *Hyman (1985: 8) points out that in some languages only a subset of the consonants make their syllable heavy when they occur in coda position*. Hayes dice que *this can be described by placing restrictions on β in the language-particular version of the Weight by Position rule*. (1989: 356). Pensamos que la idea de proyectar de abajo arriba y de filtrar por el camino puede dar cuenta de estos hechos. Habrá que investigar más sobre este aspecto tan interesante.

²⁰ Aquí se aprecia bien la diferencia entre ‘cantidad’ y ‘peso’, y se ve también la importancia de que todo el material llegue al final de la derivación, que se conserve la información a lo largo de toda la derivación.

²¹ Quizá podamos hablar de un cuarto tipo de lenguas, sensibles tanto al peso del núcleo como al de la coda. En esta categoría cabe el antiguo islandés, que diferenciaba entre *fat* /fat/, *fát* /fa:t/, *fatt* /fat:/, *fá* /fa:/, y *fátt* /fa:t:/ (ver Ewen & van der Hulst 2000: 149, en referencia a Lass 1984: 257-8).

²² Remitimos a Haegeman (1994: 412): “*Rizzi’s proposal to account for the example given here is NOT that subjacency is irrelevant for Italian. Rather he proposes that the bounding nodes are parametrized, i.e. that different languages may have different bounding nodes.*”

4 Cómo se forman las CS (II)

Hemos dedicado el capítulo anterior a definir el ámbito general de las CS en forma de espacio fonológico, y ahora llevamos nuestra atención a un punto muy concreto de ese espacio. Vamos a ver cómo se forman las CS en la confluencia entre segmentos y nodos silábicos.

Estructura general de las CS

Si bien las distintas lenguas pueden tener distintos *porqués* para formar una CS, el *cómo* debe ser fundamentalmente el mismo en todas ellas, y el inglés no será una excepción. En esencia, una CS consiste en asociar un segmento consonántico al nodo nuclear de una sílaba, o sea, una C a un N. En los términos más laxos posibles, la CS viene a ser:

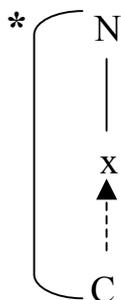
(1)

N
|
C

Pero “eso”, sin más, no debe de ser una CS en toda regla. No podemos conformarnos con definir una CS como “consonante asociada a un núcleo silábico”. Hay que profundizar en los detalles estructurales, y en esa tarea tenemos que ver en qué se salen las CS de los mecanismos habituales, dado su carácter marcado. Pero al mismo tiempo hay que ver hasta dónde se salen, porque si escaparan a los mecanismos del sistema, en vez de ‘marcadas’ serían sencillamente imposibles. ¿Cuál es el método para formar las CS, entonces?

En primer lugar, para que un segmento sea **silábico** tiene que estar asociado a la **cabeza** de N. Porque un segmento puede ser nuclear sin ser silábico, como lo es el complemento de un diptongo (la /i/ de /bai/, por ejemplo). Y también así se concibe la /-r/ de *more* o *park* en inglés (Harris 1994: 257-64; Ladefoged 2001: 58). La “nuclearidad” es necesaria pero no suficiente, pues. En segundo lugar, no debe permitirse cualquier asociación de C a N, así sin más: es necesario regular o restringir el proceso. En concreto, la inserción directa de C en el nodo N sería demasiado holgada desde el punto de vista teórico, muy sobregeneradora:

(2)



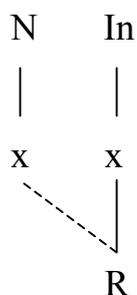
Y además esa holgura chocaría con los hechos observados, con la escasa distribución de las CS en las lenguas del mundo y aun dentro

de las propias lenguas que las tienen, como hemos visto. Partimos de la base de que lo no-marcado es asociar N sólo a una vocal, a rasgos [-consonánticos] o a elementos estrictamente vocálicos, o sea a {A, I, U, @}, según hemos argumentado ya¹.

Tenemos entonces que la C debe asociarse a la cabeza de un nodo nuclear N, y que no debe permitirse la inserción directa porque sería demasiado poco restrictiva. Con estas claves deducimos que la C está asociada en origen a otro nodo, y que desde ahí se asocia a la cabeza de N por expansión autosegmental. Así se ha propuesto en la FR, que es la única teoría que ha descendido al detalle estructural, y por eso la tomamos como referencia para nuestro estudio.

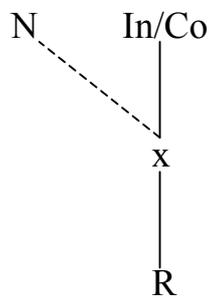
La FR propone justamente que para que la C pueda asociarse a N, este tiene que estar vacío. El concepto de **núcleos vacíos**, propio de la FR, nos parece esencial para entender la formación de CS, y con él se consigue la mejor explicación al respecto que conozcamos. La FR sostiene que en origen la C está asociada a un nodo Inicio y desde ahí se expande al N vacío colindante². Toft (2002) y Scheer (2003b) representan así este proceso de expansión autosegmental (*spreading*):

(3)



Este es el esquema básico del que partimos, y ahora lo vamos a ir modificando³. Nuestro primer cambio consiste en asociar la R a una sola unidad de tiempo, porque en caso de asociarse a dos unidades tendríamos CS largas (y bimoraicas), lo cual no es cierto⁴. Proponemos que la R siga ligada a una sola aspa de la gradilla temporal, y que la expansión autosegmental se traslade entonces al plano silábico, de esta manera:

(4)

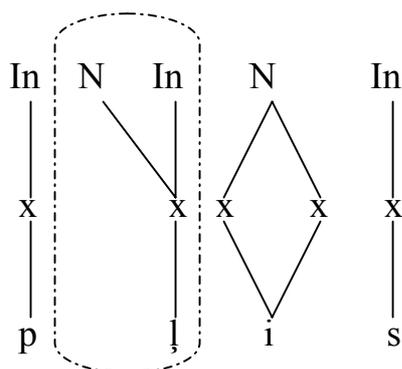


En el diagrama hemos optado por mantener la notación de la FR, que toma como posición de partida de la CS un inicio. Sin embargo, hemos anotado también que el nodo podría ser igualmente una coda. En las CS/ __# es donde hay conflicto terminológico. Para la notación tradicional, el punto de partida será una coda, y para la FR será un inicio. Sin embargo las CS/ __C tienen que estar en una coda, incluso en los términos de la FR, puesto que el inicio está ocupado por la C que sigue a la R. Nos referimos concretamente a las CS formadas tanto en ‘codas complejas’ (*thousand, important, patience* o *herald*) como en el interior de palabra (*messenger, passenger, mistletoe, Musulman*).

Por otra parte, en las CS/ __V (*analyze, platinum* o *fiddler*) la R̥ tiene que funcionar como inicio de la V que le sigue. Es evidente que la CS funciona como inicio de la V siguiente, y al mismo tiempo se asocia al núcleo vacío previo. Por consiguiente, el segmento R̥/ __V pertenece a dos sílabas a la vez, y decimos que es **ambisilábico**. Con ello estamos tomando partido a favor de la ambisilabicidad, que es un concepto controvertido. Pero francamente, no encontramos mejor forma para dar cuenta de las CS en contexto vocálico.⁵ Y más generalmente, conceptuamos todas las CS como ‘**ambinodales**’, en virtud de la estructura ‘compartida’ que hemos propuesto para representarlas⁶.

Ilustramos la ambisilabicidad con un caso concreto. La dificultad para asignar una CS/ __V a una sílaba en concreto se ve muy claramente en Mora Bonilla (2003). Este autor no acierta a representar el acento en la forma sincopada de *police* y la transcribe como [p̣li:s], sin la tilde que marca el acento de palabra. Al intentar asignar nosotros la tilde, comprobamos que resulta inadecuada en cualquiera de las dos posiciones posibles, tanto [p̣ḷi:s] como [p̣ḷi:s]. En el primero de los casos, da la impresión de que la /ḷ/ esté demasiado “separada” de la vocal siguiente. Y en la segunda transcripción parece que sea la propia CS quien reciba el acento de palabra, y además queda separada de la sílaba previa, a cuyo núcleo debe estar unida. En definitiva, da la impresión de la /ḷ/ que se oye en la pronunciación rápida o coloquial de la palabra *police* está a caballo entre dos sílabas, y así lo plasmamos en la representación gráfica:

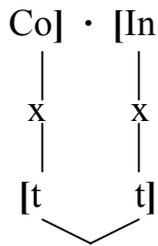
(5)



Un detalle crucial de la innovación que estamos proponiendo es que una sola unidad de tiempo sirve de origen a las dos líneas de asociación que parten hacia los nodos silábicos. Con ello hay una proporción 2:1 que recuerda a la que ya se conoce en las africadas, vocales largas y consonantes geminadas. Lo peculiar de las CS es que la **‘asociación 2:1’** tiene lugar en el plano silábico y no en el plano segmental. Esta podría ser una de las claves del carácter marcado de las CS.

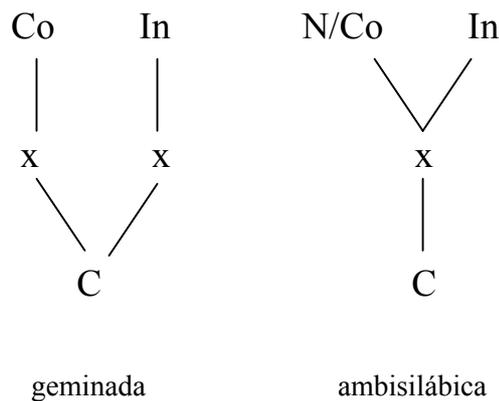
Hemos dicho además que las CS/ __V son ambisilábicas, y podríamos ampliar el alcance de nuestra propuesta para llegar a caracterizar todos los segmentos ambisilábicos en general, más allá de las CS. Si en efecto existen estos segmentos, hay que formalizar la intuición de que una C ambisilábica no es igual que una geminada. Se puede considerar a las geminadas como “ambisilábicas” en la medida en que sus “patas” se reparten entre dos sílabas, una en una coda y la otra en un inicio. Esta es la representación de una /t/ geminada⁷:

(6)



La diferencia se hace evidente a la luz de nuestro nuevo esquema. El detalle decisivo es que la geminada se asocia a dos unidades de tiempo, y la ambisilábica sólo a una:

(7)



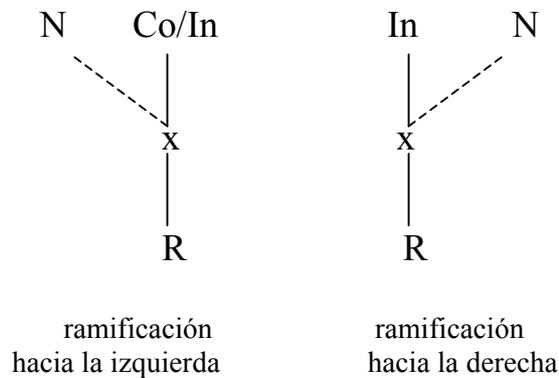
La diferencia exacta se codifica, entonces, en función de cuál sea el plano que contiene la asociación 2:1. La idea del diagrama es que las geminadas se asocian a dos unidades de tiempo (esto es, el pivote está en el plano melódico) y las ambisilábicas a dos nodos silábicos (el pivote está en los nudos del esqueleto autosegmental). Se liquidan así las dudas de Harris (2004: 4) respecto a que la ambisilabicidad pueda quedar difuminada y confundirse con la geminación, porque ahora las unidades de tiempo son distintas⁸.

Así pues, con sólo mejorar la articulación entre planos, hemos resuelto varios problemas de una vez, y sin necesidad de recurrir a nuevos aparatos teóricos. Es sorprendente cuánto de lo que ocurre está produciéndose en el plano superior. Y cómo no se aprecian los problemas de un modelo hasta que representamos todos los diagramas y nos preguntamos por todos los detalles del proceso. Otro dato de interés es que se pueden reutilizar estructuras y reglas ya existentes (la asociación 2:1 es un ejemplo) para aplicarlos en otro lugar y producir nuevos efectos. Esta combinación de simplicidad y complejidad parece ser muy propia de los organismos vivos, y un motor de su evolución y su riqueza.

Sentido de la ramificación: la Hipótesis TVR

El esquema básico de las CS reza que se forman al asociarse a un N mediante expansión autosegmental. Según si la posición de origen de la CS está en la sílaba posterior o anterior al N, hablamos de ramificación hacia la izquierda o hacia la derecha, respectivamente:

(8)



En sus primeros trabajos sobre CS, Scheer utilizó el sentido de la ramificación como criterio estructural para distinguir entre consonantes silábicas y lo que él ha denominado **consonantes atrapadas** (*trapped consonants*). Estas consonantes se diferencian de las silábicas por su falta de peso métrico y por la imposibilidad de portar el acento de palabra⁹. Hasta hace poco, Scheer ha sostenido que las CS ramifican hacia la izquierda y las C atrapadas lo hacen hacia la derecha.

Sin embargo, Scheer (2009) ha optado por asignar la ramificación hacia la derecha para las CS, en sintonía con la hipótesis de Szigetvári y Blaho (dentro de la Fonología VCVC). Y en contraste con las CS, las consonantes atrapadas ya no se asocian con núcleo alguno, lo cual se corresponde con su incapacidad para recibir el acento de palabra (que es una propiedad nuclear, según Szigetvári & Scheer 2005).

El concepto de consonantes atrapadas introducido por Scheer nos parece una contribución importante de este autor a la teoría fonológica. Por supuesto que estamos de acuerdo con la distinción fundamental entre consonantes silábicas y atrapadas, y sobre los términos establecidos por Scheer. También suscribimos la idea de que las C atrapadas no estén asociadas a ningún núcleo o nodo N, en coherencia con su comportamiento. Pero no estamos de acuerdo con la revisión del sentido de ramificación, y nos adherimos a la postura original de que las R ramifican hacia la izquierda.

Como vamos a ver, la ramificación a la izquierda es coherente con la hipótesis de que las CS son un eslabón dentro de un proceso de reducción fonológica, según hemos defendido. Recordemos que este

proceso ocurre en el plano métrico del inglés y que se pasa de vocal plena a vocal neutralizada o reducida (la schwa, en inglés), y de ahí a CS y a núcleo vacío (llamamos ‘supresión’ a este último paso):

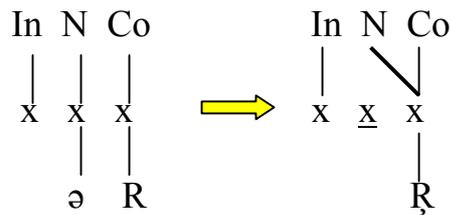
(9)



Cadena de reducción vocálica progresiva
(no tienen por qué seguirse todos los pasos de la cadena).

En este proceso de reducción progresiva, la schwa está en el núcleo previo a la R candidata a CS, y por eso postulamos que es donde se produce la asociación entre R y el núcleo:

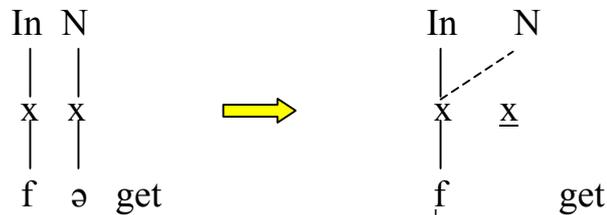
(10)



Por eso es coherente defender a la vez la ramificación hacia la izquierda y la formación de CS por reducción. Scheer ha defendido este mismo criterio durante años, pero en sus propuestas más recientes ha optado por la ramificación hacia la derecha, que relaciona con la epéntesis o inserción. La ramificación hacia la derecha se justifica por un motivo interno a la teoría de la Fonología CVCV. Se postula que hay varios núcleos vacíos y que la ramificación hacia la derecha es la única que permite licenciar el exceso de estos (para una explicación más detallada, véase Scheer 2009).

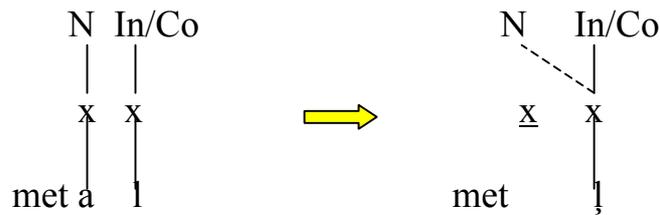
Una teoría integradora debe tener en cuenta los dos tipos distintos de CS que existen. Gracias a que en inglés la CS alterna con una forma ‘completa’ con schwa, nos es posible establecer dónde está la vocal elidida¹². Los diagramas que siguen nos muestran que en *forget* o *support* está después de la fricativa (ramificación hacia la derecha), mientras que en *cotton* o *metal* la que se elide es la vocal que va antes de la resonante (ramificación hacia la izquierda):

(12)



sin embargo,

(13)



En definitiva, el sentido de la ramificación determina la naturaleza pero no la existencia de una CS.

La génesis de CS se puede recoger de manera unificada en lo que vamos a llamar ‘**Hipótesis TVR**’. Partimos de una ‘sílabo tipo’ o ‘**sílabo canónica TVR**’: el inicio lo ocupa una obstruyente ‘T’, el

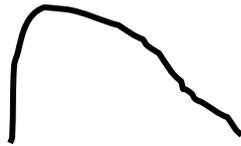
núcleo una vocal ‘V’ y la coda, una resonante ‘R’.¹³ A la hora de formar una CS, la ʈ sólo puede asociarse desde la posición prenuclear, y la ʀ sólo puede hacerlo desde la posición postnuclear. La TVR se complementa con la *RVT, que excluye las ʀ prenucleares y las ʈ postnucleares. Es este el motivo por el que en *pilot*, ninguna de las dos consonantes, ni la /t/ ni la /l/, puede formar una CS, ya que tendrían que ramificar al lado contrario al que les corresponde a cada una¹⁴. Por lo tanto, ʀ siempre ramifica a la izquierda y ʈ siempre ramifica a la derecha.

Al menos en inglés, cuando se compite por un ‘locus’ gana la R frente a la T. Si bien este hecho probablemente se deba más al índice de sonoridad que a un factor estructural, al menos la hipótesis TVR permite plasmarlo formalmente, puesto que la coda está más cerca del núcleo que el inicio. Un efecto derivado de esta competencia es que la ʈ inglesa tenderá a formarse en una sílaba TV(T) [o bien TV·T], ya que si fuera una TVR siempre saldría ganando la R. Creemos que nadie ha formulado esta observación ni esta condición hasta ahora. Todo lo más, quedaba plasmada en los ejemplos, pero se pasaba por alto. Asimismo, la cercanía estructural estipulada en la hipótesis TVR permite deducir que en todas las lenguas $ʈ \supset ʀ$.

En apoyo de la hipótesis TVR traemos a colación el **Sonority Dispersion Principle (SDP)**, de Clements 1990, que da lugar a un esquema silábico tal que una consonante es tanto mejor inicio cuanto menos sonora sea, y tanto mejor coda cuanto más sonora. Como explica Kenstowicz (1994: 283-84), el SDP produce un contorno sonoro donde el inicio tiende a producir una “ruptura” respecto a la sonoridad de la vocal, y así actúa como **baliza** para marcar el

comienzo de una nueva sílaba. Por su parte, la coda tiende a mimetizarse con el núcleo, sin romper apenas el perfil y llevando a un aterrizaje suave de la sonoridad. En definitiva, el esquema sonoro de una sílaba tiende a ser:

(14)



En un perfil que tiene todo su sentido cuando lo vemos en sucesión:

(15)



La T del inicio es como un muro que rompe con el perfil de la sílaba anterior y marca el arranque de la sílaba propia. La R de la coda, en cambio, procura no romper el contorno silábico, para quedar agrupada o integrada dentro de su sílaba. En nuestra opinión, este perfil refuerza los argumentos a favor de la existencia de sílabas. Y la asimetría entre la T del inicio y la R de la coda puede justificar la existencia de la rima como constituyente intermedio, dentro de un esquema X-barra como el que venimos utilizando en esta tesis.

Pero además, la propuesta del SDP encaja muy bien con otra idea directriz de esta tesis. Entendemos que el contorno silábico tiene como fin el procesamiento cognitivo de los elementos sonoros. Todo el

plano superior actúa como **interfaz** o mediador entre lo puramente fónico del plano melódico y los módulos cognitivos del cerebro/mente, de acuerdo con la definición del espacio fonológico que hemos trazado. En él, las sílabas organizan el material sonoro para facilitar su procesamiento cognitivo.

Y para redondear el modelo, parece que funciona bien la idea de tener un patrón característico (In (N Co)) con TVR, que se repite o itera de manera indefinida. Porque este patrón de ruptura y caída remeda el perfil de **primer y segundo plano** que ya hemos planteado para la métrica inglesa (con una sucesión de pies donde alternan sílabas tónicas y átonas). En este nuevo contexto, el inicio **destaca** o rompe más respecto al núcleo, mientras que la coda tiene un perfil más **discreto**, al estar más pegada al N. Nos encontramos pues con el patrón de **luz y sombra** o primer y segundo plano pero trasladado a componentes de menor tamaño. Es decir, la misma idea fundamental, el mismo patrón cognitivo pero expresado en dos ámbitos distintos, y esta es una de las razones de ser de esta tesis.

Cerramos pues este capítulo con una corroboración del planteamiento establecido en los capítulos anteriores. Las CS quedan establecidas en su naturaleza precisa y diferenciadas de las consonantes atrapadas, y nos quedamos con la reducción como mecanismo formador. Asimismo hemos planteado la hipótesis TVR para establecer cuál debe ser el sentido de la ramificación y lo hemos integrado dentro del esquema de luz y sombra orientado al procesamiento cognitivo. Finalmente, hemos visto cómo este mismo mecanismo opera en dos lugares distintos, una idea capital de la **Biología Evolutiva del Desarrollo** que también da sentido a nuestro trabajo.

¹ Para los aspectos teóricos de la Fonología de la Rección, remitimos al Apéndice I de esta tesis.

² Solamente Toft propone la inserción directa de /l/ en el nodo nuclear, que ya hemos rechazado por su excesivo poder generativo. En cambio, Rennison & Neubarth (2003: 10) dicen que el elemento fonológico **R** no puede estar asociado léxicamente al núcleo: cuando **R** está asociado a un N (en vocales rotificadas y en coroneales silábicas) es siempre por expansión desde O.

³ Nosotros preferimos conservar las unidades de tiempo, pero Scheer ya no. En sus trabajos recientes, desde la ‘Fonología CV Estricta’, prescinde de la gradilla de tiempo y asocia los segmentos del plano melódico con nodos C y V, directamente. Esos nodos vocálicos y consonánticos sirven como esqueleto autosegmental, de manera análoga al esqueleto CV ya propuesto antes de las unidades de tiempo, pero siempre con una cadencia estricta CVCV, es decir, alternando siempre nodos C y nodos V.

⁴ Los estudios acústicos recientes y aun el sentido común desmienten que la CS sea larga; Yu (1999) indica que las CS más bien se abrevian, y en inglés del Estuario (*Estuary English*) se cambia la /l/ por una /ʊ/, según Cruttenden (2001: 87-88); este hecho se aprecia incluso cuando la /l/ es silábica (por ejemplo, *middle* puede pronunciarse /ˈmɪdʊ/ (2001: 88), y no hay duda de que es una vocal corta.

⁵ Por supuesto, las CS del checo no deben de ser ambisilábicas, puesto que no existen CS/___V en esta lengua.

⁶ Al estar asociada a dos posiciones distintas, una CS tiene dos valores contradictorios, aunque no incompatibles. Tiene a la vez carácter silábico (por su asociación con la cabeza del núcleo silábico) y no-silábico (por asociarse a un inicio o una coda). No los consignamos en forma de rasgo, pero sí que aprovechamos para recordar que no es lo mismo ‘nuclear’ que ‘silábico’: este último sólo marca la posición de cabeza del N.

⁷ Como recoge el diagrama, la /-t/ de coda tiende a ser implosiva, mientras que la /t-/ de inicio es aspirada, o al menos debe ser más fuerte que la de una coda, ya que los segmentos varían ligeramente en función del constituyente subsilábico que ocupen.

⁸ Kessler (1998) comenta que las lenguas se caracterizan por tener o bien geminadas o bien ambisilábicas o bien ninguna de ellas, pero lo que no hay es lenguas que tengan a la vez geminadas y ambisilábicas. Sin embargo, Harris (2004: 107) dice que “*this is demonstrably false*”, y lo argumenta diciendo que “*Ibibio has short intervocalic consonants that would qualify for coda capture in any ambisyllabist’s book, but it also has authentic geminates in the same context*”.

Por otra parte, queremos hacer notar la diferencia de nuestra propuesta con la que da la Fonología CVCV, que prescinde de las unidades de tiempo en sus diagramas (aunque también es capaz de representar la diferencia entre consonantes geminadas y silábicas).

⁹ Scheer (2007) reconoce como consonante atrapada a cualquier **cúmulo extrasilábico** de inicio de palabra. El español no tiene consonantes atrapadas, por ejemplo, pero sí que aparecen en lenguas tan diversas como el bereber, árabe marroquí, griego, lenguas eslavas, lenguas caucásicas y salish. Scheer (2007b) da como ejemplos en árabe marroquí *rbīl*, *rDa*, *lga*, *kbīr*, *zna*, *ndīm*, *qbīl* (‘ligar’, ‘aceptar’, ‘encontrar’, ‘llorar’, ‘cometer adulterio’, ‘lamentar’, ‘aceptar’). Obsérvese que, a diferencia de las geminadas parciales, las C atrapadas no concuerdan ni en el rasgo [sonoro] ni en el de [LUGAR] con la C siguiente.

A diferencia de las CS, las C atrapadas no tienen especificaciones sobre su naturaleza: una vez que una lengua admite consonantes atrapadas, las admite todas. Basta con un parámetro que marque un umbral para decidir si una lengua permite o no estas consonantes. Es la condición ‘todo vale’ (*anything goes*) de Clements (1990). Otra cuestión es cuánto se realice en la práctica; pero lo cierto es que siempre permiten que se amplíe el repertorio, por ejemplo a la hora de tomar préstamos.

Al igual que las CS, las C atrapadas pueden ramificar tanto a izquierda como a derecha. La mayoría ramifican hacia la izquierda, y sólo las del polaco, que están en la rama derecha de un inicio, ramifican hacia la derecha. Scheer (2003b) cita *trwać* ‘durar’, *drzwi* ‘puerta’, *grzmieć* ‘tronar’ o ‘fulminar’, *brzmieć* ‘sonido’ o ‘leer’, *trzcina* ‘caña’, y *klnę* ‘lo juro’, *plwocina* ‘esputo’. Es difícil llamar ‘coda’ a esta estructura peculiar, puesto que a la derecha de ese ‘inicio ramificado’ tiene que haber un N vacío; por ejemplo, *grz0mieć*.

Al igual que nos tenemos que plantear cuál es el origen de las CS, tenemos que hacerlo con las consonantes atrapadas. La falta de peso métrico que las caracteriza nos hace pensar que se forman a partir de un inicio. Es de suponer que para formar las C atrapadas se extienda una rama desde el inicio inicial de palabra. La rama sale hacia la izquierda y se forma una especie de coda, con la peculiaridad de que el núcleo que precede a esa coda no está lleno, sino vacío (y de ahí su falta de peso, una vez más). Este mecanismo probablemente no sea exclusivo; también podrían formarse por expansión de un inicio hacia la sílaba previa las geminadas, la nasalización de las oclusivas en terena y el *raddoppiamento sintattico* en italiano.

¿Cómo se legitima entonces esta ‘coda’ de nuevo cuño? De acuerdo con la FR se hace una legitimación ‘horizontal’ desde el inicio a la coda. Una segunda posibilidad es que se haga mediante una asociación 1:2, que probablemente sea común para los demás tipos de expansión hacia la izquierda de un inicio. Y bien podría ser la propia asociación 1:2 (combinada con la especificidad del lugar donde ocurre, quizá) la que estuviera codificada en el parámetro que distingue a las lenguas con C atrapadas.

La tercera conjetura es que tengan algo en común con las consonantes en anacrusa (como la s en *stay*, la ʃ en *schmuck* o la f en *phthisis*, que seguramente sean consonantes atrapadas), y que entonces estén sujetas a una forma de ‘magic licensing’ (una propuesta de Kaye, dentro de la FR). Por otra parte, una consonante como la /f/ de *FNAC*, que no tiene un sentido de ramificación claro, podría tener algo de “flotante” (hemos oído pronunciar [efnac] y [fenac]). Y una cuarta posibilidad es que las C atrapadas se adjunten al nodo de sílaba o de palabra, usando un mecanismo ya propuesto en la teoría fonológica.

¹⁰ Recordemos los ejemplos del capítulo 1 y el estudio de Ridouane (2002).

¹¹ Obsérvese que en estos casos la ʈ es una fricativa: el rasgo [+continuo] parece importante.

¹² Para reforzar esta argumentación podemos “revertir el proceso” y probar a insertar vocales. Ya hemos visto cómo la vocal insertada va antes de la sonante R en el uso expresivo de *great*, *please*, *freak out* (oído en un episodio de *South Park*). Por el contrario, si queremos insertar una vocal en las palabras que contienen CS en tashlhiyt imdlawn, creemos que nadie dudaría en colocar la vocal después de la ʈ. Retomando ejemplos que citamos en el capítulo 1, *sɣf* ‘desvanecerse’ tiende a ser *sɣɔf* más que *səɣf*, pero sobre todo, *sʃqqsɔt* ‘irritarle’ apunta claramente hacia *sʃəqqəst*, en nuestra modesta opinión.

¹³ Estamos hablando en los términos tradicionales de inicio y coda, pero incluso en la FR la coda es siempre una resonante R (el inicio puede ser una R o una T). La sílaba TVR es en esencia una sílaba cerrada (por mucho que también haya CS / __V). Esto implica que todas las R tienen algo de coda en su naturaleza. Por supuesto, las CS / __V tienen también algo de inicio, como ya hemos visto.

¹⁴ Harris (1994: 185) da una lista de palabras que son susceptibles de síncope. La lista es larga, y entre las palabras se incluyen *fact(o)ry*, *choc(o)late*, *cam(e)ra*, *pris(o)ner*, *fam(i)ly* o *comp(a)ny*. Todas ellas cumplen con el requisito TVR, como vemos. En cambio, en la lista siguiente vienen otras palabras en las que la síncope es cuando menos dudosa, o más bien imposible. Estas palabras

son *rock(e)ting*, *mon(i)tor*, *goss(i)ping*, *ball(o)ting*, *opac(i)ty*, y *men(a)cing*, y en todas ellas se ha invertido el perfil sonoro de TVR, esto es, son RVT, y por eso no es posible la elisión. Por lo tanto, la hipótesis TVR parece estar haciendo predicciones acertadas.

5 Cómo se forman las CS (III)

Nuestro análisis sobre el modo en que se forman las CS (el *cómo*) ha ido descendiendo en el grado de detalle. El recorrido comenzó en lo ‘macroscópico’ (desde su aspecto más general, el de la articulación del espacio fonológico en planos) y se ha dirigido gradualmente hacia lo ‘microscópico’, que es el detalle más fino. De este modo hemos alcanzado el nivel segmental, que es el que nos disponemos a examinar en este capítulo.

Vamos a profundizar en los detalles estructurales de la formación de CS con la atención puesta en la zona de articulación entre segmentos y nodos silábicos. En este nivel de la estructura, pretendemos que la formación de CS quede encuadrada entre los procesos de reducción, y que lo haga sin que sea necesario recurrir a ningún mecanismo extraordinario, sino solamente haciendo uso de los mecanismos ya existentes en el sistema.

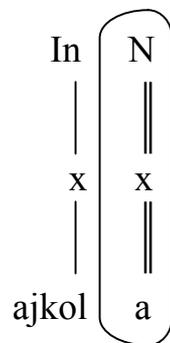
Procesos de reducción parcial (1):

Venimos asumiendo a lo largo de esta tesis que las CS se forman por reducción, mediante la regla $\partial R \rightarrow R$, que es una regla de nivel segmental. Ahora es el momento de estudiar en detalle cómo funciona

este mecanismo. A tal fin vamos a sugerir un análisis conjunto de varios procesos de reducción, y para ello seguiremos el modelo holístico y de organización espacial concreta que venimos defendiendo. Recordemos que en este modelo el *cómo* es también el *dónde*. Con este modo de articular el espacio fonológico se puede obtener una diversidad de resultados a partir de una sola operación de deleción: para ello basta con aplicar esa única operación en lugares distintos de la estructura. Vamos a ver cómo.

La forma más extrema de reducción es aquella que implica la pérdida de toda una posición fonológica: los segmentos, las aspas y los nodos silábicos. Este proceso es el que hemos denominado ‘**supresión**’, de la cual encontramos un ejemplo ilustrativo en la pérdida de las vocales finales en chukchi (Spencer, 1996: 65):

(1)

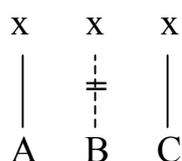


El resto de los procesos de reducción no son tan extremos, y siempre conservan algo. Pero incluso para la supresión nos parece que hay un componente de conservación; bien mirado, las líneas de asociación no se pierden, y por eso las hemos marcado dobles en el diagrama. Y así es como la pérdida de un segmento acarrea la pérdida de todo el resto de la posición, a la cual arrastra consigo.

Con esta perspectiva en mente, vamos a asumir que todos los procesos de reducción combinan la eliminación con la conservación, en una tensión armoniosa que se atiene a lo que Ewen & van der Hulst (2000: 169) han denominado **efectos de estabilidad**. Nuestra postura es que la eliminación nunca es total ni indiscriminada, sino que siempre se conserva algo, incluso en los casos más extremos. Es más, nosotros pensamos que se puede establecer un gradiente de eliminación, que empieza por la supresión, sigue con el alargamiento compensatorio (A-C), y termina con la coalescencia.

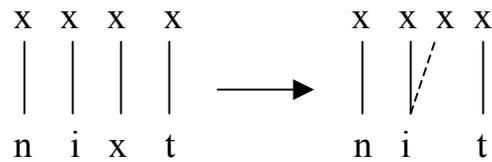
El **alargamiento compensatorio (A-C)** también se caracteriza por la eliminación de un segmento, pero a diferencia de la supresión, el segmento se desliga de su posición en la estructura, y por lo tanto no se conservan las líneas de asociación inferiores:

(2)



En el A-C, la disociación de un segmento supone la reasociación de otro de los segmentos, que ve aumentada su duración. Es lo que sucede en latín **kasnus* → *ka:nus*, y **fideslia* → *fide:lia* (véase Hayes 1989 para una explicación en detalle), o en inglés *nixt* → *ni:t*. El formalismo autosegmental representa el proceso como:

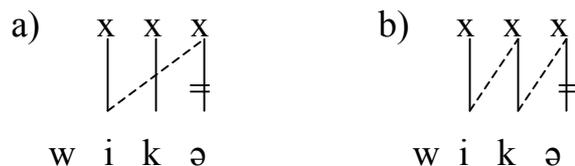
(3)



Ahora bien, la explicación autosegmental clásica nos parece menos satisfactoria cuando los constituyentes no tienen adyacencia inmediata. Es lo que vemos en el **MEOSL (Middle English Open Syllable Lengthening)**, el proceso en el que /wikə/ > [we:k(ə)] ‘week’, así como /wudə/ > [wo:d(ə)] ‘wood’ y /nosə/ > [nɔ:s(ə)] ‘nose’ (véase Ewen & van der Hulst 2000: 172). Vemos que en el MEOSL la pérdida de una schwa final acarrea el alargamiento de la vocal previa, pero entonces la schwa y la vocal están en sílabas distintas, y además hay segmentos (consonánticos) que se interponen.

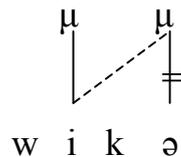
Pues bien, para representar el paso de *wikə* a *we:k* nos veríamos obligados a cruzar líneas, en flagrante violación de un principio autosegmental¹ básico (solución ‘a’); o como alternativa, podríamos hacer un desplazamiento de las aspas del esqueleto para reasignar a los segmentos afectados (solución ‘b’, análoga a la presentada en 2000: 172).

(4)



Ninguna de las dos soluciones puede ser satisfactoria, evidentemente. Los escollos de este modelo quedan salvados en la fonología moraica, que ignora el segmento interpuesto y permite que las moras queden estrictamente adyacentes:

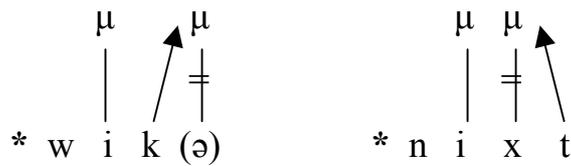
(5)



Ahora no hay segmentos que se interpongan: crucialmente, el MEOSL se hace en sílabas abiertas, por lo que las consonantes no tienen valor moraico y no cuentan para la conservación de las moras, y entonces no interfieren en el proceso. De este modo, el alargamiento compensatorio no se limita a una conservación de la cantidad (en las aspas autosegmentales), sino que implica también una conservación del peso (en las moras).

Pero nosotros no terminamos de quedarnos contentos con esta explicación. No se nos escapa que al eliminarse la schwa en *wikə*, la consonante /k/ pasa a ocupar una posición de coda, y con ello debe adquirir valor moraico. Esta reasignación de las moras debería bastar para conservar el peso métrico, pero no se ha tenido en cuenta esta posibilidad. Lo mismo podemos decir de *nixt*, porque cuando se elimina la /x/, la /t/ pasa a tener valor moraico, pero se ignora esta solución también:

(6)

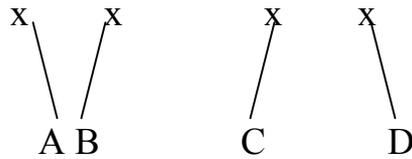


La solución pasa por conservar el peso métrico (las moras) a costa de los segmentos ya licenciados como moraicos previamente, en el aducto (*input*) de la regla. Pero estipular una condición así sin más es demasiado arbitrario y carece de poder explicativo. ¿Qué es lo que está ocurriendo aquí?

Nuestro modelo considera tanto el esqueleto autosegmental como la gradilla moraica. Es el modelo de estructura integrada que venimos defendiendo a lo largo de toda esta tesis, y que ahora con el estudio del A-C (alargamiento compensatorio) se pone a prueba en otro ámbito. Vamos a comenzar con el ejemplo más sencillo, el de *nixt* → *ni:t*. Volvemos la atención hacia el plano segmental y la gradilla de tiempo, como hacía el estudio autosegmental. Pero aquí vamos a introducir un cambio de perspectiva.

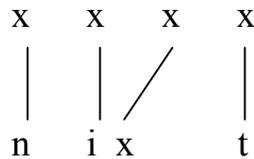
El modelo autosegmental clásico consistía en una serie de procesos de asociación y disociación, tomando el segmento como pivote de las líneas de asociación hacia la gradilla de tiempo. Pero en nuestra propuesta el pivote se invierte: ahora está en las aspas del esqueleto autosegmental, y son los segmentos los que se desplazan. Gracias a este nuevo pivote, los segmentos pueden acercarse o alejarse entre sí, un hecho que representamos con ‘**ángulos de enlace**’:

(7)



Los segmentos A y B se acercan entre sí, en tanto que los segmentos C y D se alejan. Es de notar que las aspas que hay entre A y B están a la misma distancia que las que hay entre C y D. ¿Qué queremos decir con que dos segmentos se ‘acercan’? En el caso de *nixt*, el acercamiento de los segmentos /i/ y /x/ se traduce en un proceso de asimilación gradual, que termina con la plena identificación de la /x/ con la /i/. En concreto, la /x/ puede lenificarse primero a /h/, después a /j/, y finalmente a /i/. Hay cada vez mayor afinidad entre ellos:

(8)



Los ángulos de enlace reflejan una resonancia creciente que se está produciendo en el nivel infrasegmental, y que a su vez va a repercutir sobre toda la estructura. O sea que la aproximación está sucediendo realmente en los rasgos infrasegmentales; los rasgos de la /x/ se van modificando progresivamente para acercarse cada vez más a los de la /i/, en un proceso análogo al de la evolución histórica, que pasó por una lenificación de la fricativa (/niht/).

Así es como /x/ se acerca a /i/, dejando entrever que se va a producir el A-C. El movimiento no nace en la /i/, sino en la /x/, que se va

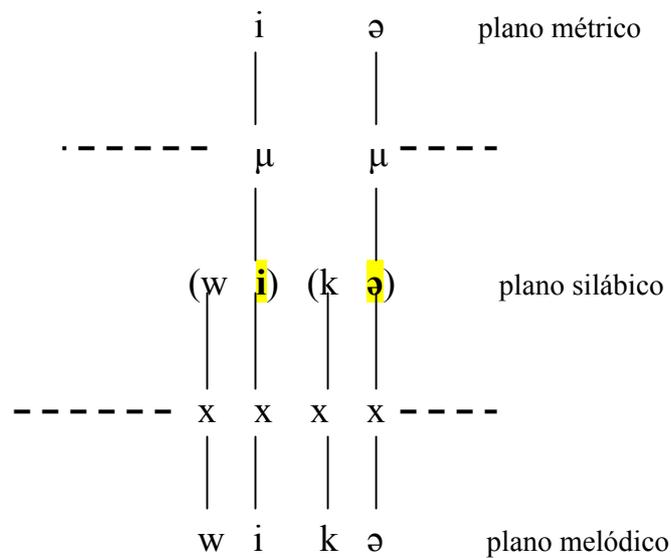
acercando cada vez más a la /i/ hasta que finalmente se funde con ella. El resultado final sigue siendo una /i/ con dos líneas de asociación al esqueleto autosegmental, pero la manera en que se ha llegado a esta situación es significativamente distinta.

Llega el turno de analizar el MEOSL, con las diferencias y similitudes que pueda tener con *nixt*. Es evidente que tenemos que salir del plano melódico si no queremos incurrir en las contradicciones ya señaladas. Y visto el éxito de la fonología moraica para superarlas, vamos a desplazarnos hasta el plano superior, que es donde hemos situado las moras (recordando la máxima de que el *cómo* es también el *dónde*). Es más, nosotros vamos a postular que el A-C tiene su propio ámbito: dada la coincidencia entre MEOSL y palabras trocaicas, pensamos que aquel se realiza dentro de un troqueo².

Otra innovación respecto a la explicación moraica tradicional es la manera precisa de proyectar hacia niveles más altos. Para este caso nos decantamos por probar con la **estructura de frase escueta** (*Bare Phrase Structure*) del **Programa Minimista** que presentamos en el capítulo de introducción. Esto implica que la proyección a niveles más altos se realiza con los propios segmentos; los que son cabeza de su constituyente ascienden a niveles más altos de la estructura, mientras que los complementos se quedan abajo.

En las palabras del MEOSL, son justamente las vocales las que reciben el peso moraico (en virtud de su ‘peso por posición’), y por eso ascienden, en tanto que las consonantes se quedan abajo:

(9)



[N.B.: en el plano silábico hemos marcado en negrita los segmentos que funcionan como núcleo (N) de la sílaba. Dado que el MEOSL se produce en palabras con sílabas abiertas, las consonantes tienen sólo valor de inicio y por tanto no van a recibir peso métrico. Así pues, en el nivel más alto las vocales quedan en contacto directo, sin que se interponga ninguna consonante.]

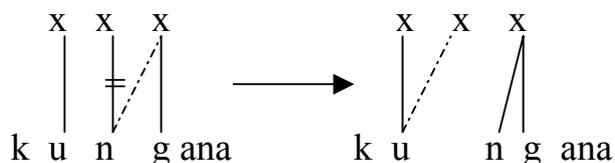
Al igual que en *nixt*, los dos segmentos que han ascendido al plano métrico entran en **resonancia** (aunque esta vez se trata de dos vocales). La /ə/ se acerca gradualmente a la vocal principal del troqueo, la /i/ en este caso, hasta que finalmente converge con ella. Digamos de paso que es el complemento el que se acerca a la cabeza para converger, y que esta norma también parece regir la convergencia de segmentos en *nixt*. Lo único que parece diferenciar los dos fenómenos es que en el MEOSL (pero acaso también en *nixt*) la acción se traslada al plano métrico. En definitiva, la regla es la misma, pero opera o se expresa en un lugar diferente de la estructura³.

Nuestra propuesta tiene varias ventajas teóricas, que sepamos. Por una parte, salva el escollo de los constituyentes discontinuos, al igual que hacía la fonología moraica, pero es que además se evita la reasignación de moras a las consonantes, y al tiempo establecemos la reacción entre dos segmentos que ya son moraicos. Por mucho que se invocara todavía la conservación del peso, el modelo integrado haría que el mecanismo ya no sea igual.

Y la posibilidad de operar con la misma regla en distintos lugares de la estructura sirve a la perspectiva unificada que buscamos para tratar los procesos de reducción en general. Sólo con una perspectiva integrada se puede dar cuenta del mecanismo que opera en una serie de procesos en apariencia dispares. Y gracias a este enfoque estamos en condiciones de concluir que el sistema fonológico se vale de mecanismos relativamente simples para sus computaciones, y que basta con aplicar un único mecanismo (que es sencillo) en lugares diferentes del armazón fonológico para obtener una diversidad de efectos muy productiva⁴.

En apoyo del razonamiento de que el alargamiento compensatorio no es una forma pura de eliminación, aportamos un ejemplo de la lengua bantú kiñaruanda que nos parece aclaratorio. Se trata de un tipo de A-C peculiar, donde el A-C sirve para compensar la reasignación de segmentos, pero crucialmente no se produce ninguna eliminación de material melódico (véase Roca 1994: 10-11). Por ejemplo, en la palabra *kungana* hay que resilabear la primera /n/ para mantener abiertas todas las sílabas (kun·ga·na → ku·nga·na), y entonces se alarga la /u/ para conservar el peso inicial (ku:ngana):

(10)



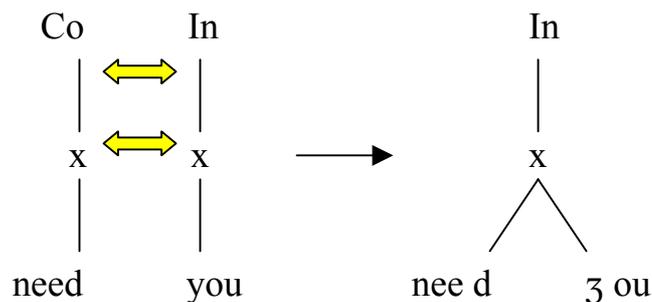
El diagrama tiene dos detalles importantes, que ponen de relieve la naturaleza gradual de los procesos de reducción. Uno es que se ha producido igualmente un A-C para conservar el peso y la cantidad, en los mismos términos que ya conocemos⁵, pero sin eliminar material melódico, que simplemente se ha reasignado. Así es como damos un nuevo un paso en el gradiente de eliminación decreciente que estamos siguiendo.

El otro detalle es que la consonante reasignada pasa a formar un cúmulo /ng/ en el que hay dos consonantes ligadas a una sola aspa autosegmental. Esta es la estructura típica de las africadas, que se forman precisamente por coalescencia. Por lo tanto, el ejemplo de la reasignación de segmentos en kiñaruanda queda a medio camino entre el A-C y la coalescencia. Si trasladamos este esquema al ejemplo de *nixt*, podemos entender que cuando la distancia entre /i/ y /x/ llegue a ser cero, habrá una fusión de dos segmentos por proximidad. Este es el fenómeno que se conoce como **coalescencia**, y nos parece que el A-C no está tan lejos de ella.

Reducción parcial (2): la coalescencia

La afinidad entre A-C y coalescencia no es nueva. El propio Spencer (1996: 201-2) considera como A-C la coalescencia que se produce en la **palatalización de coronales** del inglés. Es un fenómeno muy común, y podemos verlo en *what you want*, *would you*, *in case you* y *has your* (1996: 201),⁶ cuando la /j/ forma una sibilante palatal por asimilación con la obstruyente previa ([wɒtʃu:wɒnt], [wʊddʒu:], [ɪŋkeɪʃʃu] y [hæʒʒɔ:], respectivamente, de acuerdo con la notación de Spencer). Estos casos representan un paso más allá en el equilibrio entre pérdida y conservación. Lo que se conserva es el material melódico, pero al tiempo se “pierden” componentes del plano silábico o del esqueleto. Las líneas pueden fundirse o replegarse, por lo que la eliminación de un componente arrastra a los otros componentes que estaban asociados mediante esas líneas:

(11)

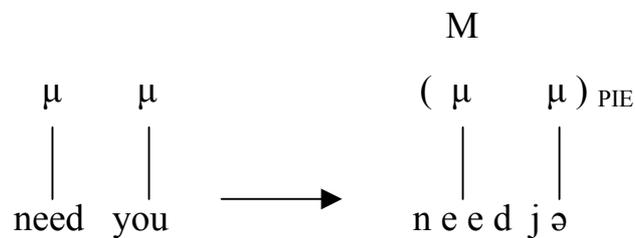


Se ha formado así una africada, que tiene una estructura 2:1, concretamente, “dos segmentos/un aspa de tiempo”; esto implica una reducción parcial, que acorta la duración temporal de los segmentos. Dado que estas africadas no son subyacentes, la estructura se alcanza a partir de dos posiciones independientes, como refleja el diagrama. Y

aunque en este primer intento lo hemos plasmado como un proceso en bloque, nos da la impresión de que en realidad consta de varios subprocesos.

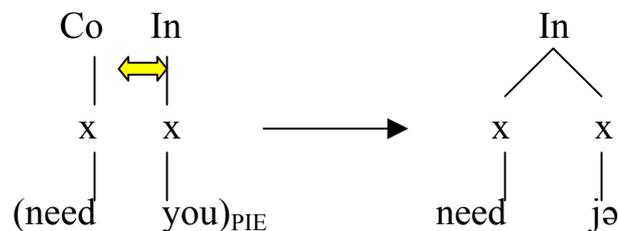
El primero es la necesidad de agrupar *need* y *you* en un solo pie métrico, que es el troqueo {*need you*} (¡otra vez la métrica en inglés!). Porque es en este contexto donde sobreviene el debilitamiento o neutralización de la vocal de *you* a schwa:

(12)



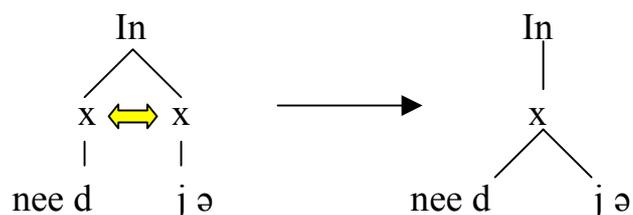
En el plano silábico, los nodos terminales de *need* y *you* se van acercando (¡otra vez un acercamiento, como el que postulábamos para el A-C con los ángulos de enlace!), hasta el punto de confluir en un solo inicio ramificado /dj-/:

(13)



A su vez, esta ramificación del inicio debe transferirse al esqueleto fonológico, con el resultado de que las dos aspas del diagrama se funden en una sola:

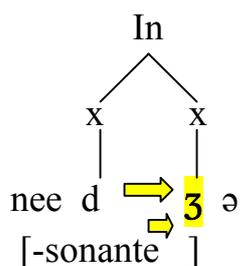
(14)



Y al fusionarse la unidad de tiempo, también le siguen las líneas de asociación entre inicio y aspa, que se repliegan para formar una sola, según vemos.

En lo que respecta al plano segmental, el valor del rasgo [-sonante] se extiende⁷, gracias a una regla autosegmental, desde la /d/ de *need* hasta la /j/ de *you*. El resultado es que la ‘yod’ /j/ se hace fricativa, /ʒ/, es la misma resonancia que apuntábamos en *nixt*, y que es en realidad infrasegmental:

(15)

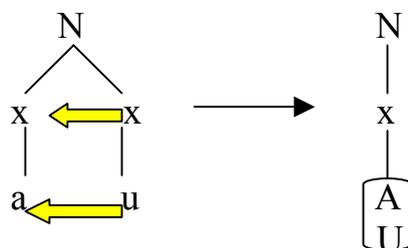


Estos procesos pueden hacerse de arriba abajo o de abajo arriba⁸. Podemos inclinarnos por lo primero, teniendo en cuenta lo que ocurre en otro proceso de coalescencia de segmentos: nos referimos al que reduce los diptongos a monoptongos, por ejemplo, en [ai → e] y en [au → o]⁹. En este caso partimos ya de una posición acaso más

favorable para la coalescencia, porque los segmentos ya están agrupados bajo un mismo nodo silábico, que es N. Por consiguiente, sólo queda que confluyan las dos aspas del esqueleto y que las líneas de asociación hagan lo propio en el plano silábico.

Cuando miramos en el plano melódico, la absorción de diptongos difiere ligeramente de la formación de africadas, porque la relación entre segmentos parece ser más simétrica. En lugar de transferir un rasgo de un segmento a otro, da la sensación de que los dos segmentos se funden en uno solo, y sus rasgos o elementos confluyen (en este sentido, nos parece especialmente gráfica la explicación que da la fonología de partículas (Schane 1995: 586-608), y por extensión, la de la fonología de elementos (Harris 1994: 114-16)):

(16)



Coalescencia para formar monoptongos

Nótese aquí el punto de asimetría que se obtiene al postular a una de las vocales como cabeza de la estructura. Se colige que el acercamiento corre a cargo de la otra vocal, que es el complemento, y de ahí que las flechas sean unidireccionales. Es más, el esquema de **cabeza y complemento** se repite a lo largo de todos los ejemplos que llevamos vistos. En *wike* era (**i** ə), en *nixt* era (**i** x), en *need you* era (**d** j), y ahora en *au* es (**a** u). El mismo patrón estructural, con la cabeza a la izquierda y el complemento que se le aproxima.

En conclusión, hemos argumentado que podemos dar cuenta de una diversidad de procesos a través de un solo mecanismo. Se reduce o se agrupa en un lugar determinado de la estructura, pero al tiempo se conserva todo lo demás. Existe una tensión entre eliminar y conservar, y la clave es que no se elimina todo: hay que decidir qué se elimina y qué se conserva. El mecanismo es productivo, dado que con una sola operación obtenemos tres fenómenos distintos: la supresión, el alargamiento compensatorio y la coalescencia. Los mecanismos operativos son sencillos, y la complejidad se obtiene por la complejidad en la organización del espacio fonológico, y gracias a la posibilidad de operar en distintos puntos o ámbitos de ese espacio. El acercamiento gradual (representado mediante ángulos de enlace) y la estructura de cabeza y complemento son detalles que nos aportan más luz sobre la reducción.

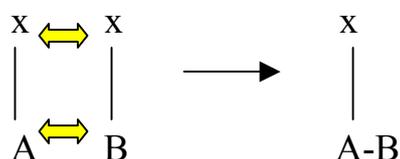
Formación de CS por coalescencia

Partiendo del mecanismo general de reducción parcial que acabamos de plantear, el siguiente paso es encuadrar la formación de CS dentro de este mecanismo. Se trata de entender la formación de CS como un fenómeno de eliminación parcial cualquiera, como uno más, que no se diferencia esencialmente de los que llevamos vistos. La ventaja evidente de esta solución es que las CS se forman utilizando mecanismos ya existentes, y así quedan perfectamente integradas en el aparato fonológico. O sea que no hace falta postular nada nuevo, porque ya existe: las CS son solamente un proceso más de entre todos los que ya hay. En definitiva, por mucho que las CS sean un proceso

‘marcado’ o excepcional, ese carácter especial que tienen no es debido a los mecanismos estructurales utilizados para su formación.

Aún podemos llevar más lejos la analogía, y asimilar la formación de CS a una suerte de coalescencia. Como en los ejemplos que hemos estudiado, también aquí hay una confluencia de las aspas del esqueleto autosegmental acompañada de la fusión del componente segmental:

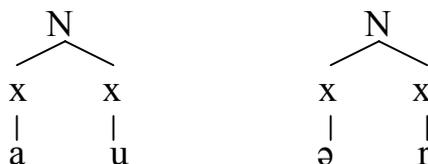
(17)



La fusión del material melódico (segmentos y rasgos) da cuenta de la reducción de material que caracteriza a las CS, y la fusión de las aspas da cuenta de su corta duración (como un solo segmento) y de su menor peso métrico (que se corresponde con un menor peso cognitivo). Y dado que es un proceso de reducción parcial, hay también una parte que se conserva: los nodos del plano silábico. Otra característica común de coalescencia y formación de CS es la posibilidad de revertir el proceso¹⁰; sólo así puede obtenerse la oscilación entre /əR/ y /R/ propia de las CS débilmente silábicas.

Vamos a trasladar el esquema de la coalescencia a la formación de CS, pues. Para ello comenzamos por la formación de /r/ silábica, por el gran parecido estructural que apreciamos con la coalescencia de diptongos. En efecto, tanto los diptongos como la /-ər/ están alojados bajo la cobertura de un nodo nuclear N:

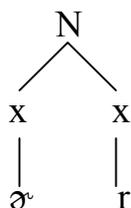
(18)



La teoría de que en inglés la /r/ está alojada en el nodo nuclear N, que nosotros asumimos en esta tesis (Harris 1994: 257-64), se fundamenta en la capacidad que tiene esta consonante para “teñir” las vocales del núcleo. Por ejemplo, las vocales /e, i, u/ quedan neutralizadas en una sola vocal /ɜ/ cuando están en presencia de una /-r/ (*perfect*, *fir* y *fur* son todas /ɜ:/). Y según constata Ladefoged (2001: 58), también les ocurre algo similar a otras vocales: “ (...) *in most forms of American English in words such as ‘snarl’* [snarl], *when /r/ has to be considered part of the vowel.*”

Pues bien, la schwa también resulta afectada por la /r/ en el cúmulo /-ɚr/, hasta el punto de que en inglés se habla de una **r-colored schwa** y se la representa con un símbolo especial, /ɚr/. Su estructura fonológica es así:

(19)



El ‘coloreado’ de la vocal denota que esta absorbe características de la /r/ o se “tiñe” de sus rasgos. De modo que podemos verla como una suerte de coalescencia, tanto desde el punto de vista conceptual como estructural. Es más, podemos equiparar la ‘r-colored schwa’ a la coalescencia de diptongos que se produce dentro del núcleo silábico, en la medida en que obtenemos $[\vartheta+r \rightarrow \text{r}]$ por el mismo procedimiento que $[a + i \rightarrow e]$ y $[a + u \rightarrow o]$:

(20)



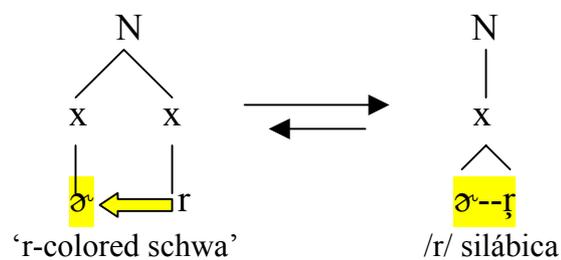
donde no es necesario retocar los rótulos de nodo para establecer la coalescencia, sino que basta con fundir las aspas del esqueleto y también fundir los segmentos correspondientes. Las dos figuras del diagrama ponen de manifiesto el paralelismo de las dos formas de coalescencia: la /r/ en la rama derecha de un nodo N es análoga a la de la semiconsonante de un diptongo. Este hecho no es de extrañar, dado que la /r/ del inglés comparte con las paravocales su articulación como aproximante.

En el aspecto estructural, no está de más recordar aquí que la fusión no es del todo simétrica. Recordemos que para los procesos de reducción hemos postulado una relación de cabeza y complemento, con la cabeza en el lado izquierdo del constituyente ramificado. Para las CS se cumple el que la /ə/ rige a la /r/, pero sorprende cómo es la consonante quien tiñe a la vocal de sus rasgos, y deberá ser la

consonante quien proyecte a niveles más altos. En este aspecto podemos estar ante una cierta anomalía estructural, donde el complemento parece ser el que se proyecta, y no la cabeza¹¹. De ser así, estaríamos ante una de las claves del carácter marcado o excepcional de las CS.

Sea como sea, de momento nos mantenemos en el modelo de fusión que convierte a la ‘r-colored schwa’ en una /r/ silábica:

(21)



En este diagrama hemos plasmado cómo las aspas del esqueleto han confluído y la ramificación del nodo N se ha trasladado al esqueleto autosegmental. La flecha de retorno en el proceso indica la oscilación entre /ə/ y /r/, para dar cuenta de la forma débilmente silábica /-(ə)R/, con un grado variable de proporción de schwa y de consonante.

El estudio acústico de Yu (1999) sobre sibilantes en chino refuerza esta hipótesis en dos sentidos. El primero es que la comparación del cantonés *si* ‘seda’ con su equivalente mandarín *sz* demuestra que la duración total de las dos palabras es igual, y lo que varía es la proporción de los componentes (1999: figura 1). En la del mandarín *sz* se ha alargado el inicio /s-/ porque la silábica es significativamente más corta. Con este mismo patrón, asignamos una sola unidad de

tiempo (*timing unit*) a la CS. El segundo motivo es que la sibilante silábica tiene componentes acústicos a la vez vocálicos y consonánticos. En cambio, el inicio es sólo consonántico, tanto en *si* como en *sz*, mientras que la /i/ de *si* es únicamente vocálica.

Por eso nos parece adecuado incluir la /ə/ y la /r/ juntas en la representación fonológica de la CS. Los resultados de Yu casan muy bien con el hecho de que las CS tienden a parecerse mucho a la vocal del núcleo, a ser muy resonantes, mientras que el inicio del contexto izquierdo tiende justamente a marcar distancia, a diferenciarse en cuanto a sonoridad y a otros rasgos (así lo veremos en el capítulo 8, al hablar del contexto izquierdo).

La verdad es que nunca se ha dejado del todo claro cuál es la diferencia acústica entre lo que se ha dado en llamar ‘r-colored schwa’ y la [ɹ], y no sería de extrañar que la diferencia fuera sólo de grado o incluso inexistente. Esta sospecha parece confirmarse con el estudio de Yu, donde la formación de una sibilante silábica no parece que haga desaparecer del todo la vocal; más bien, es como si la vocal se hubiera teñido del sonido de la consonante. Es decir, que igual que la candidata a CS se acerca a la posición de la vocal, también la vocal parece que se acerca a la consonante.

Como colofón, la fusión total de schwa y /r/ produce la silábica /ɹ/, que representamos según el esquema general de cualquier CS, si acaso con la salvedad de que sólo hay un nodo silábico, que es N:

(22)

N
|
x
|
r

Una vez establecido el mecanismo de formación de la /r/ silábica, el siguiente paso es trasladar este razonamiento a su consonante hermana, la /l/, con la cual constituye la clase natural de las líquidas. De nuevo postulamos una coalescencia como la de los diptongos, sólo que esta vez la /l/ está alojada en la coda y no en el núcleo silábico. A pesar de esta diferencia, se conservan una serie de paralelismos entre las líquidas y las paravocales en inglés que en líneas generales hacen viable la hipótesis de la coalescencia. Vamos a ver cuáles son estos paralelismos.

El primero es el grado de sonoridad: líquidas y paravocales están en la gama intermedia de la escala sonora, en la frontera entre vocales y consonantes. Abstrayendo la división cualitativa entre unas y otras, vemos que paravocales y líquidas están contiguas en la escala numérica de sonoridad¹². El segundo es que en la especificación de rasgos las líquidas son [+consonante] [+sil], mientras que las paravocales son el reverso de estos rasgos, porque son [-consonante] [-sil].¹³ Una y otra asociación de rasgos rompen con el esquema más canónico que asocia [α consonante] con [β sil]¹⁴. El tercero es que pueden ocupar la rama derecha del inicio (aunque no hay *tl-/dl- en inglés). Las paravocales tienen la peculiaridad de ser además las únicas vocales que pueden estar en el nodo In (incluso en el nodo no

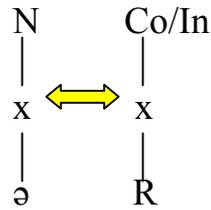
ramificado), igual que la /r/ es la única consonante que está en N (pero sólo ramificado).

Las paravocales y la /r/ comparten un cuarto paralelismo, y es que ambas pueden ocupar la rama derecha del nodo N, como complemento del núcleo (por ejemplo, *day*, *bow*, *part*). Aquí no está incluida la /l/, pero nos parece una diferencia menos significativa, que no basta para descartar la hipótesis de la coalescencia. Al fin y al cabo, la diferencia de nodos silábicos no era impedimento para formar las africadas /tʃ-/ y /dʒ-/ en *I want you/need you by my side*.

En definitiva, los tres primeros paralelismos deben bastar para hacer extensivo el esquema de la coalescencia a la otra consonante líquida, la /l/, y seguramente al resto de las sonantes (es decir, a las nasales). Como es natural, habrá que tener en cuenta que esta vez la coalescencia no es “intra-nodo”, dentro del nodo N, sino que afecta a nodos silábicos distintos. Pero no es menos cierto que la propia /r/ puede sufrir coalescencia sin estar en el nodo N: en concreto la r / __ V (por ejemplo *several*, *opera* o *wondering*) ocupa el inicio de la sílaba siguiente, y no la rama derecha del nodo N. Es más, una CS de este tipo tiene carácter ambisilábico (ver *infra*, capítulo 10), y por tanto es fundamental que los rótulos de nodo se conserven intactos, sin confluir en uno solo (a diferencia de lo que ocurría en la confluencia de inicios que vimos para la formación de africadas).

Así es como se forma cualquier CS por coalescencia. Partimos de un núcleo ocupado por una /ə/ y una coda (o un inicio) ocupada por una ‘R’. La operación se realiza cuando se acercan entre sí (representado por la flecha doble) las dos aspas del esqueleto autosegmental:

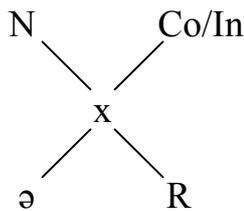
(23)



‘Primer paso’ de la formación de CS

El acercamiento gradual concluye con la fusión de las dos aspas en una sola, como corresponde a un patrón de coalescencia:

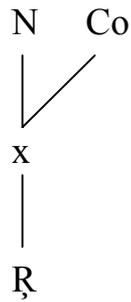
(24)



‘Segundo paso’ de la formación de CS

Con este esquema tenemos un primer esbozo de CS. Hemos conservado los dos segmentos melódicos ligados a una sola unidad de tiempo (como en las africadas), porque así damos cuenta de las CS débilmente silábicas, donde se permite la coexistencia de vocal /ə/ y sonante ‘R’, con grados variables de realización de cada una de ellas. En las CS fuertemente silábicas se pierde la vocal por completo (por ejemplo, en *mountain* /^lmaʊnʔŋ/) y solamente queda la CS, en estado puro y ocupando toda la posición:

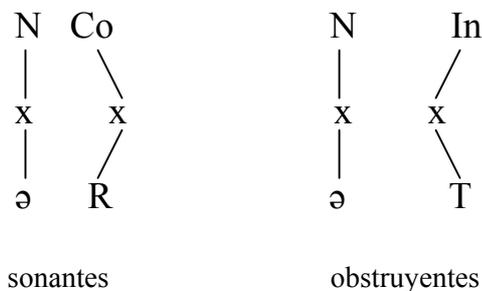
(25)



El desarrollo del razonamiento sobre la coalescencia ha vuelto a dejar claro que unas consonantes son más idóneas para formar CS que otras. La más apta es la /r/, puesto que la coalescencia se realiza dentro de un mismo nodo silábico (el N) y que la /r/ cumple con los cuatro paralelismos con las paravocales. La /l/ es algo menos apta, pero mantiene buenas condiciones, y otras consonantes van resultando menos aptas a su vez.

Nosotros sugerimos que los distintos grados de aptitud de las consonantes queden plasmados en la representación fonológica. A tal fin, proponemos que la distancia sonora se traduzca en la correspondiente distancia estructural, con lo cual habrá diferentes ángulos de enlace entre los componentes de la sílaba. Por ejemplo:

(26)



Así, las consonantes estarán fonéticamente tanto más cerca de la vocal nuclear cuanto más resonantes sean, de manera que el grado de sonoridad de las consonantes determina su ubicación en la sílaba¹⁵. Obsérvese que las aspas del esqueleto están exactamente igual de separadas en los dos casos, y lo que cambia es el ángulo de enlace (igual que en los enlaces químicos) de las líneas de asociación. Esta es una manera novedosa de tratar las líneas de asociación, que siempre han sido verticales, con lo que el modelo era más estático y no recogía estos matices.

El último obstáculo que debe salvar el nuevo modelo teórico es evitar la sobregeneración. Es necesario establecer condiciones en el esquema general de las CS, de manera que la expansión hacia la rama derecha de un N esté permitida para ciertas consonantes pero vedada para otras. Una solución sencilla es establecer una **distancia crítica** entre la sonoridad de la schwa¹⁶ y la de la candidata a CS. La restricción reza que para que una consonante llegue hasta la rama derecha de N, debe estar ya muy cerca, no más allá de cierta distancia crítica¹⁷; superada esta, la coalescencia resulta inviable. La distancia crítica probablemente sea paramétrica, y lo habitual es que sólo las líquidas y si acaso las nasales queden dentro de su alcance.

Observación final

El estudio de la coalescencia debe entenderse dentro de un enfoque más amplio. Los matices introducidos demuestran que no sólo cuenta la suma total de componentes. Si únicamente nos fijáramos en el balance estequiométrico, la coalescencia podría reducirse a una

pérdida de material en la trama estructural: pérdida de un nodo In en el primer paso de la formación de africadas, y pérdida de una unidad de tiempo (*timing unit*) en el segundo paso. Sin embargo, creemos que la idea de fusión de dos componentes iguales en uno solo es más fiel a lo que ocurre.

Otra ventaja que le vemos a este primer esbozo es el buen correlato estructural que tiene con la idea de resonancia. La disposición de las líneas de asociación en diagonal (y no en vertical), y formando entre todas una cruz, nos transmite la sensación de resonancia con claridad.

Finalmente, la autonomía que aún conservan los constituyentes en esta disposición permite **volver atrás** en el proceso. En otras palabras, el esquema capta la idea de que la formación de CS es **reversible**, y por eso puede alternarse entre la forma no-silábica /-əR/ y la silábica /-R/. Y además de permitir la reversibilidad, esta estructura también permite los distintos grados de realización de la /ə/ en la forma fonética, desde la vocal perfectamente pronunciada hasta el cero, pasando por todo un espectro de formas intermedias. Daremos más detalles en el capítulo 7.

¹ Goldsmith (1976), en Goldsmith (1999: 146), da dos condiciones de buena formación (diagrama 22: *Well-formedness conditions (initial statement)*), la segunda de las cuales es “*Association lines do not cross*”.

² Goldsmith (1989: 73-6) analiza el alargamiento compensatorio en turco, a partir de Sezer (1985). A nosotros nos parece que un análisis ‘moraico’ da cuenta de los datos perfectamente. Es posible que en turco el A-C no se asocie a un troqueo, pero sí a la rima (y por tanto las moras). Por ejemplo, a raíz de una delección opcional de consonante, hay A-C en *savmak* → *sa:mak* ‘librarse

de', pero no en *savul* → *saul* 'tambor' (**sa:ul*). También lo hay en *sahte* → *sa:te* 'falsificación' o *seyret* → *se:ret* 'ver', y no lo hay en *tohum* → *toum* 'semilla' o *deyil* → *deil* 'no es'.

En apoyo del análisis que hemos planteado, podemos decir que el A-C de estos ejemplos es propio del habla informal o relajada, cuando la consonante se hace más laxa, y resulta más propensa a la coalescencia.

³ Este fenómeno recuerda al de la expresión de genes en los seres vivos. Según el segmento del cuerpo en el que nos halleemos, se da lugar a una estructura anatómica determinada o a otra. Por ejemplo, García Bellido explicó cómo la expresión del gen *Ultrabithorax* en el tercer segmento torácico de la mosca induce la formación de halterios (pequeños órganos de balanceo), y su falta de expresión en el segundo segmento produce alas. Cuando el gen regulador deja de expresarse en el tercer segmento, se producen alas aquí también (con lo cual la mosca no tiene halterios y tiene dos pares de alas), en tanto que si se expresa en el segundo segmento, se producen halterios en esa zona (y la mosca no tiene alas pero tiene dos pares de halterios). Véase Sampredo (2002: 148 ss.) para una explicación muy instructiva.

⁴ Recordemos la productividad que tiene la fonología autosegmental con sólo usar asociaciones y disociaciones.

⁵ De hecho, Roca ni siquiera considera esta posibilidad, pero es bien posible que de nuevo se trate de un fenómeno métrico y no puramente segmental.

⁶ Otro ejemplo de formación de africadas es el que cita Spencer (1996: 67) para el sufijo comparativo /-ʃi:/ en checo, donde /t + ʃ/ o /d + ʃ/ se simplifica a [tʃ]. Por ejemplo, *star-i:* 'viejo' da *star-i:* 'más viejo', pero *tlust-i:* 'gordo' da *tlustf-i:* 'más gordo', y *mlad-i:* 'joven' da *mlatf-i:* 'más joven' [Spencer 1996: 67].

⁷ O bien el elemento ? en la FR.

⁸ Una cosa es la construcción de la estructura, que se hace siempre de abajo arriba, como hemos establecido en la introducción, y otra cosa es cómo se realizan los procesos de ajuste o cambio. En la coalescencia que estamos estudiando, no basta con la interacción primaria entre los segmentos, porque lo primero es entrar en el ámbito de un troqueo.

⁹ Hay tres maneras de reducir los diptongos, al menos en inglés británico. Una es el **diphthong smoothing** (Cruttenden 2001: 139-40), que supone perder la paravocal del complemento, de la rama derecha. Por ejemplo, (aɪə → aə) en *fire*, (eɪə → eə) en *player*, (ɔɪə → ɔə) en *employer*; y (əʊə → ə:) en *slower*, o (aʊə → a:ə) en *shower*. O incluso, (a:ə → a:) y (a:ə → a:) en formas refinadas de RP (*Received Pronunciation*) (2001: 139).

Otra es perder la cabeza del diptongo y mantener la paravocal, que es lo que ocurre en *Monday* ['mʌndɪ] o en *always* ['ɔlwɪz], así como en *fortunate* o *chocolate* y *necklace*, *preface*, *village*, *manage* (2001: 106-8).

La tercera es la coalescencia, donde no se pierde material melódico pero sí se agrupa este bajo una sola unidad de tiempo, según se explica en la Fonología de Partículas (véase Schane 1995: 591) y se adopta en la Fonología de Elementos (véase Harris 1994: 98-9). En inglés, Harris cita ay → ε: (*bait*) y aw → ɔ: (*caught*) (1994: 99). En los tres casos, la reducción implica pasar de dos unidades de tiempo a una sola.

¹⁰ El proceso puede revertirse en inglés, pero no en checo, según veremos en el capítulo 12.

¹¹ En la nasalización de vocales también vemos cómo la influencia va de derecha a izquierda, del complemento a la cabeza. Esta influencia parece admisible mientras se limite a un 'teñido' de la vocal, esto es, una transferencia parcial de los rasgos. Pero la vocal siempre permanece ahí como cabeza de la estructura; la consonante nasal puede permanecer o perderse.

¹² Lo relevante aquí es la sonoridad, más que los rasgos. Sólo las vocales altas pueden ser paravocales, porque son las más cerradas, las menos vocálicas, y por ello son las menos resonantes. Pero si nos fijamos simplemente en el rasgo [+alto] o [-bajo] quizá no se capte la intuición adecuadamente. Por el mismo patrón, las CS más idóneas son aproximantes, su articulación queda en la frontera entre vocales y consonantes, pero al hablar de [+consonante] y los rasgos que queramos especificar (aparte de [sonante]) creemos que se nos escapa de nuevo la intuición básica.

En definitiva, creemos que los rasgos que hay establecidos no terminan de captar la analogía entre paravocales y consonantes aproximantes (o semiconsonantes y semivocales, si queremos llamarlas así), en tanto que la escala numérica de sonoridad sí que lo hace, al menos en parte, y quizá pueda reformularse la articulación de segmentos en términos numéricos, de distancia respecto al techo de la boca en el caso de las vocales, y de grado de fuerza o presión de la lengua contra el techo para las consonantes. (Recordemos la cita de Ladefoged (2001) que dimos en el capítulo 1).

¹³ El rasgo [silábico] está en franco desuso, pero nos permitimos la licencia de usarlo aquí para marcar el carácter silábico o no de un segmento.

¹⁴ Si los nodos están rotulados con “V” y “C”, las paravocales son segmentos especificados como [-consonante] pero insertados en un nodo etiquetado como “C”, esto es, directamente en un nodo [-silábico] y sólo ahí.

¹⁵ La disposición de los segmentos en la sílaba en función de su sonoridad está en consonancia con el PSS (Principio de Secuenciación Sonora). También es una forma de articular la fonética con la fonología.

¹⁶ También la sonoridad de la schwa influye: es la menos sonora de todas las vocales, y probablemente esa sonoridad reducida sea la que permite la expansión de la CS. En línea con Yu (1999), es entonces cuando la consonante más cercana tiende a compensar el hueco de la vocal pequeña o nula. Cualquier otra vocal resultaría demasiado sonora para permitir la expansión.

¹⁷ También influye la sonoridad de la vocal que encabeza el N. Sólo cuando la vocal es muy poco resonante, hasta el punto de ser apenas perceptible, se dan las condiciones para que la consonante se acerque más al centro del núcleo.



6 Cómo se forman las CS (IV)

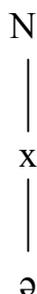
Hemos llegado al nivel más bajo en la formación de las CS. Por fin estamos en disposición de examinar los mecanismos íntimos, los detalles infrasegmentales de la coalescencia, y ver en qué nivel de la estructura se realizan. En el capítulo siguiente daremos fin a la exposición con un ascenso desde lo más pequeño a lo más grande, como forma de completar nuestra comprensión de todo el proceso.

El locus de la CS

Uno de los temas centrales de esta tesis ha venido siendo que las CS se forman por reducción de /əR/ a /R/. De aquí se deduce que el locus de la CS sólo puede ser un N ocupado por una **schwa**. La implicación deja lugar para que haya schwas sin que haya CS, como ocurre en francés. Pero si no hay schwas no hay CS, y así se cumple en lenguas como el italiano, el quechua o el vasco. Otra predicción asociada a las schwas es que no haya CS largas en origen¹, dado que la schwa es siempre una vocal corta², que ocupa una sola unidad de tiempo.

Pero no sólo es esto, sino que además la schwa debe ser la **cabeza** de su N, ya que esa es la posición **silábica**. Tiene que estar sola en el N, o bien ocupar la rama izquierda si el nodo está ramificado. En esencia, la estructura del N candidato a formar una CS es:

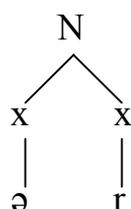
(1)



A partir de esta estructura básica desarrollamos el resto del razonamiento. Primera cuestión: para que la schwa encabece su N, debe ser (necesariamente) la única vocal del nodo. La presencia de cualquier otra vocal desplazaría de inmediato a la schwa de la cabeza³. Por ejemplo, el diptongo que hay en *there* /'ðeər/ tiene a la schwa en posición de complemento, y desde ahí no puede dar lugar a una CS (*/ðeɹ/).

En el N candidato no hay sitio para otra vocal que la schwa, pero en cambio sí que hay sitio para una consonante. Vemos que la rama derecha del núcleo ramificado /-ər/ permite alojar una /r/, sin que esta consonante compita por el estatus de cabeza:

(2)



El diagrama nos muestra un elegante correlato estructural en la más básica o primaria de las CS que conocemos, la /r/, puesto que la schwa y la /r/ figuran como **hermanas**. Ahora bien, mucho nos tememos que ni la estructura ni la coalescencia (manifestada aquí en forma de rotización) basten para definir una CS. Vamos a ver los detalles.

Es un hecho que no se forma CS en *car* o *bird*, donde la /-r/ es nuclear pero no silábica (compárese con la CS de *better*). Y esto es así a pesar de que los núcleos con {V + r} (donde V es una vocal plena) cumplen con los requisitos de cercanía estructural y de coalescencia. En efecto, la V y la /r/ ocupan las dos ramas del N, en posición de hermanas, y además la /r/ tiene algo de coalescencia con la vocal (que se tiñe de sus rasgos, especialmente la /ɜ/). ¿Por qué no se llega entonces a la silabicidad? El problema de la vocal plena, a nuestro juicio, es que no deja que la consonante se expanda sobre la cabeza del núcleo. Con lo cual, la coalescencia sólo alcanza a ser parcial, y la R no puede adquirir silabicidad. Inferimos así que la coalescencia es condición necesaria pero no suficiente para formar CS.

Hay que llegar a la coalescencia total, y para ello es imprescindible partir de una schwa. Esta vocal es la única de todo el repertorio que puede llegar a la máxima coalescencia con una consonante (candidata a silábica), que a su vez es la condición para que esta adquiera silabicidad. La schwa es una vocal tan apta para este fin que puede lograr la CS incluso en condiciones estructurales menos favorables⁴.

La idoneidad de la schwa para formar CS tiene algo que ver con su baja sonoridad. Es muy probable que sea la menos resonante de todas las vocales (y la menos vocálica), y por eso no puede competir con

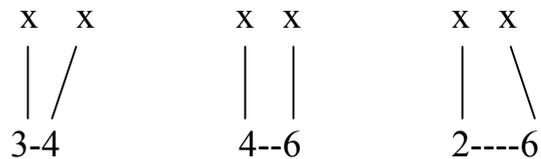
ninguna otra vocal por la cabeza del nodo N⁵. Si volvemos sobre la escala de sonoridad del capítulo 1, podemos probar a asignar un índice sonoro para la schwa. Recordemos que el mayor valor de la escala correspondía a las vocales bajas, con un índice 6, y que las vocales altas, algo menos resonantes, tenían un valor de 5. Para la schwa es bien posible que el valor sea más bajo aún: le vamos a asignar un rango que está en $[5 \geq \text{schwa} > 4]$ (el 4 marca ya a las consonantes más resonantes, y por eso ponemos a la schwa por encima de ese valor).

Como ya sabemos, el bajo índice sonoro permite una menor distancia con la candidata a CS. Para el conglomerado {ə-r}, esa distancia pudiera estar por debajo de la unidad ($\Delta\text{SON} < 1$), dado que la /r/ es la más resonante de las consonantes (especialmente la [ɹ])⁶. Pero por mucho que la distancia sonora sea determinante, no debemos perder de vista que el minuendo tiene valor fijo. Sucede así en la coalescencia de diptongos, donde no es sólo cuestión de distancia sonora, sino que los valores de los segmentos implicados se limitan a 5 y 4, sin ninguna otra posibilidad (en inglés). Y siguiendo esta misma línea, en la coalescencia de CS la schwa es el único origen posible de la resta (con su valor de entre 4 y 5). Y a partir de aquí, ya contamos con que es mejor que el sustraendo (la consonante candidata a CS) sea lo más parecido posible, siempre que la distancia no sea cero.

La distancia sonora (ΔSON) encuentra buena expresión en los ángulos de enlace que indican el acercamiento o alejamiento entre segmentos. Cuanto menor sea la diferencia sonora entre segmentos, tanto mayor será la convergencia entre ellos, y más se acercarán sus respectivos

ángulos de enlace. Y también al contrario, una mayor distancia sonora quedará reflejada en forma de alejamiento:

(3)



Los ángulos de enlace traducen la cercanía sonora a cercanía estructural. Los segmentos que están más cerca entre sí son los que después pueden entrar en un proceso de coalescencia, ya que la cercanía estructural implica cercanía entre rasgos, afinidad entre ellos. No hay coalescencia si no hay cercanía estructural.

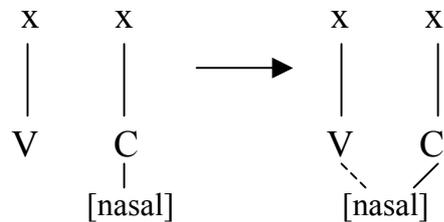
Pero ya hemos establecido que la cercanía sonora y estructural, e incluso la coalescencia (cuando es sólo parcial, como en *car*) no define la formación de una CS. Los segmentos se acercan progresivamente, pero finalmente tienen que fundirse en grado suficiente para dar lugar a una CS (y además, el proceso tiene que ser reversible, o al menos debe poder alternar de un recorrido de la derivación al otro). La clave de todo ello está en los rasgos constituyentes de los segmentos. Estamos a punto de adentrarnos en el corazón mismo del proceso.

Coalescencia de rasgos

La fonología autosegmental explica que para asimilar un rasgo se procede a compartirlo. Resulta bastante sencillo no tener que

compartir más que un solo rasgo. Pensemos en la nasalización de una vocal, por ejemplo:

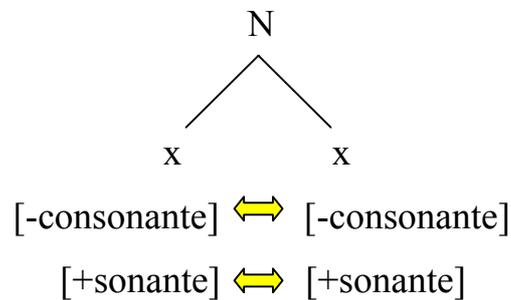
(4)



La coalescencia también entraña rasgos compartidos, a nuestro entender, partiendo en principio de los rasgos en común y de los más afines, que hemos expresado mediante la cercanía estructural (los ángulos de enlace). Pero esta vez son todo un conjunto de rasgos los que se comparten, y por eso la coalescencia es más compleja. La forma más sencilla de coalescencia es seguramente la de los diptongos, dado que ambos segmentos son de la misma naturaleza, son dos vocales⁷.

Los estudios de Fonología de Partículas han dado cuenta de esta fusión de manera muy interesante para nuestro criterio (en el capítulo anterior hablamos de Schane (1995)). Y es que a pesar de utilizar otro tipo de constituyentes (las partículas, en vez de los rasgos), nos abre el camino para traducir este mismo proceso a la fonología de rasgos. Empezamos por asumir que los nodos de nivel más alto son iguales, por lo que es fácil fusionarlos, y que el resto se reduce a operaciones con rasgos de nivel más bajo, más periféricos.

(5)



La cuestión se complica empero para las africadas, porque la naturaleza de los segmentos no es tan afín como en la diptongación, con toda probabilidad. Hay una consonante oclusiva seguida de una aproximante, que al ser vocoide está a caballo entre las vocales y las consonantes. Pero además, es que la semiconsonante se convierte en una fricativa sibilante, por lo que el proceso no se limitaría a compartir los rasgos que ya se tenían, sino que habría que introducir un nuevo rasgo, [+sibilante] o [+estridente].⁸

En el siguiente paso del gradiente, las CS, la distinta naturaleza de los segmentos a fusionar eleva la complicación al máximo. Tenemos una consonante más una vocal. La discrepancia se manifiesta en unos nodos superiores diferentes, por lo que (a priori) es de esperar que la fusión o coalescencia sea más difícil, menos probable.

En sintonía con este razonamiento, nos parece que el uso de rasgos distintos (al menos en parte) para vocales y consonantes reflejaría mejor la dificultad para compartir dichos rasgos. Una cosa es la fusión parcial (como la “roticidad” de la vocal en *bird*), pero esa total aglutinación de consonante y vocal que exhiben las CS debe ser

insólita. Y es bueno que sea así, a nuestro juicio, porque si no entraríamos en una situación de ‘todo vale’ que ni se observa ni nos parece deseable.

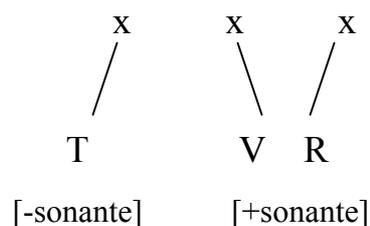
Hemos apuntado que la schwa es la (única) vocal capaz de someterse a este proceso tan peculiar. Su baja sonoridad nos habla de su escaso carácter vocálico, y aquí tiene que residir su idoneidad para entrar en coalescencia con una consonante (que a su vez tiene que ser lo más resonante posible). En inglés, la schwa es una vocal especial, porque aparte de ser neutra, se forma en la parte alta de la estructura, en el plano métrico, a raíz de una reducción vocálica. Se empieza por una vocal plena ‘V’, después una vocal reducida ‘v’, y se puede acabar en la supresión ‘0’, quizá pasando por la CS entre medias:

(6)



¿Cómo opera este proceso en el nivel infrasegmental? Comenzamos con la ‘V’: tenemos una rima formada por /-VR/, donde los dos segmentos comparten el rasgo [+sonante], que a su vez los distancia del valor negativo que tiene en el inicio T (siempre dentro de la hipótesis TVR):

(7)



En el siguiente paso, que es la reducción a schwa, se produce una mayor aproximación. En el nivel infrasegmental, esto se traduce en una asignación de valores negativos para todos sus rasgos en general. En primer lugar, destacamos el valor [-tenso], al cual atribuimos buena parte del carácter reducido y neutro de la schwa.

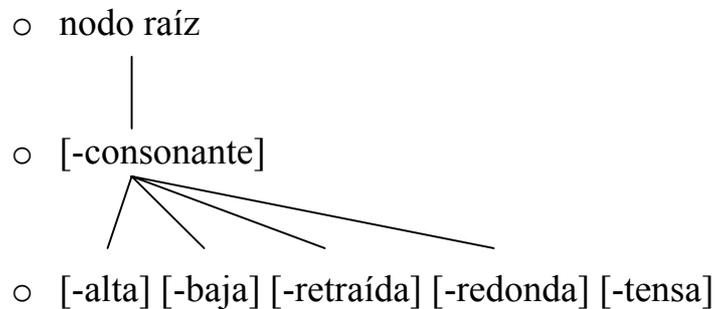
‘Reducido’, porque en inglés las vocales laxas son intrínsecamente cortas, en tanto que las vocales tensas son las que son largas. En principio, la especificación de rasgos infrasegmentales no debería decir nada sobre el número de unidades de tiempo al que se van a ligar, pero es un hecho que {ɪ, ʊ, ə, ʌ, æ} son siempre cortas. Por esto es por lo que asociamos el valor [-tenso] al proceso de reducción vocálica.

Y ‘neutro’, porque las vocales reducidas del inglés {ɪ, ʊ, ə}, son todas más centrales que sus variantes plenas⁹. Nosotros pensamos que los rasgos de posición ([alto], [bajo], [retraído]) determinan la ubicación de una vocal en el espacio fonológico, pero no son capaces de reflejar un efecto de **centralización relativa**, como el que va de /i/ a /ɪ/ o el que va de /u/ a /ʊ/. Por eso proponemos que se matice la posición general con el rasgo [±tenso], donde el valor negativo del rasgo indique una pérdida de **intensidad** en la posición. Por ejemplo, si una vocal es ‘alta’, su variante laxa será ‘alta pero menos’, y lo mismo podemos decir con respecto a otras posiciones¹⁰.

Además de esto, la posición central de la schwa en el espacio fonológico se refleja en su especificación negativa para todos los rasgos de posición propios de las vocales. En efecto, la schwa es

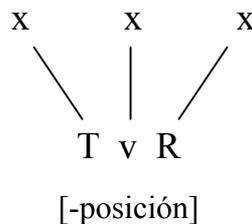
[-alta] y [-baja], por ser una vocal media, y también es [-retraída], con lo cual es además [-redondeada].¹¹ Semejante cantidad de rasgos negativos casa bien con una neutralización en círculos concéntricos, que hace que las vocales se centralicen. Podemos concluir con confianza que el paso de vocal plena a vocal neutra, de V a v, no sería más que un ir trocando a negativo los valores de todos sus rasgos inferiores¹²:

(8)



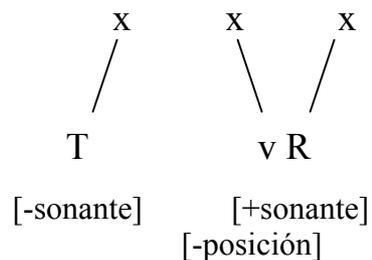
La especificación negativa de rasgos de posición de la schwa concuerda con la de las consonantes, al tiempo que marca distancia con otras vocales¹³. Si lo resumimos todo en forma de [-posición], caeremos en la cuenta de lo cerca que está la schwa de las consonantes a este respecto:

(9)



Nótese aquí que la vocal es una ‘v’ (vocal reducida), y que la coincidencia incumbe también a la T del inicio. Pero sobre todo, existe un mayor acercamiento entre N y Co, porque el número de rasgos que comparten la vocal reducida ‘v’ y la R es mayor que con la vocal plena ‘V’. Por tanto:

(10)

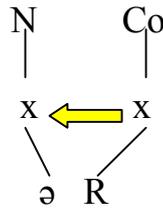


La schwa está muy cerca de las consonantes en todo lo que atañe a los rasgos de más bajo nivel, y marca distancia con estas en los rasgos más altos. Para distinguirse de las sonantes ‘R’, la schwa sólo dispone de un rasgo tan general como [-consonante]. Constatamos así que schwa y R son segmentos muy cercanos, y vamos a postular que aún se puede hacer algo para acercarlos más. La rima /-əR/ es el origen de la CS pero no es la CS misma. El paso decisivo que da carta de naturaleza a estos segmentos será el tema de la siguiente sección.

Las CS y la estructura de núcleo vacío

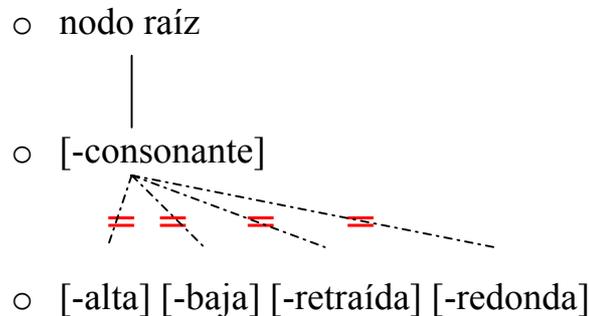
Ya que la CS sólo tiene una unidad de tiempo, y la rima con schwa tiene dos, tendremos una parte del camino andada con la reducción o la fusión en el esqueleto autosegmental:

(11)



Pero este paso fundamental debe ir acoplado a la coalescencia entre schwa y R en el nivel infrasegmental. Queremos mantener este paso dentro del proceso gradual de reducción, pero al tiempo hay que encontrar una característica que responda de la peculiaridad de las CS. Proponemos que se haga desligando los rasgos de bajo nivel de la schwa¹⁴; los rasgos de posición, desde luego, y tal vez el rasgo [+sonante]:

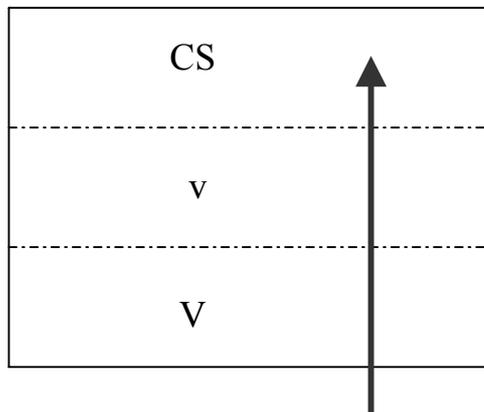
(12)



Un desligado tan masivo de rasgos tiene sus riesgos. Al fin y al cabo, no se puede ir añadiendo o quitando rasgos a discreción, porque de nuevo entraríamos en una política de ‘todo vale’ muy dudosa. No se puede recurrir al ligado o desligado de rasgos autosegmentales sin más, sólo para cubrir una necesidad de explicación: hay que justificarla, y más cuando afecta a tantos rasgos de una vez. El proceso tiene que estar bien restringido.

Hay dos factores que favorecen esta maniobra tan ambiciosa. En primer lugar, se nos antoja que la especificación negativa de todos los rasgos hace más fácil prescindir de ellos. El segundo factor está en el proceso de reducción progresiva. Es de la mayor importancia contar con que el segmento asciende a lo largo de la estructura de planos, y a medida que lo hace se van filtrando rasgos. Por eso la regla $\{V \rightarrow v \rightarrow CS \rightarrow 0\}$ debe formularse reflejando la situación de cada componente en los planos:

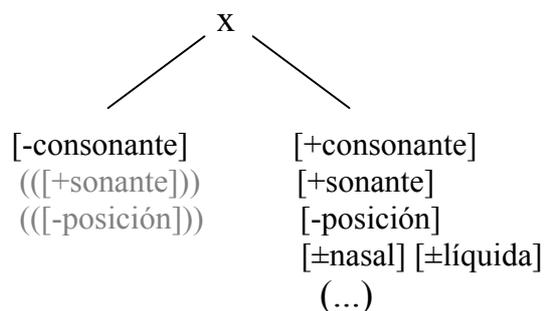
(13)



En esta estructura ascendente, los segmentos que están más abajo tienen todos los rasgos (la vocal plena), y la **reducción** significa que se **filtran** menos rasgos positivos, debido a las características de los planos más altos, y por las exigencias de otros componentes del cerebro (distintos del fonológico). Dentro de todo este esquema, podemos entender que la CS (o el desligado de rasgos de la schwa) significa que esos rasgos no tienen realización fonética, que no se van a transferir a ese componente.

Volviendo al desligado de rasgos de la schwa (que marcamos en gris y con paréntesis en el diagrama), el resultado global es:

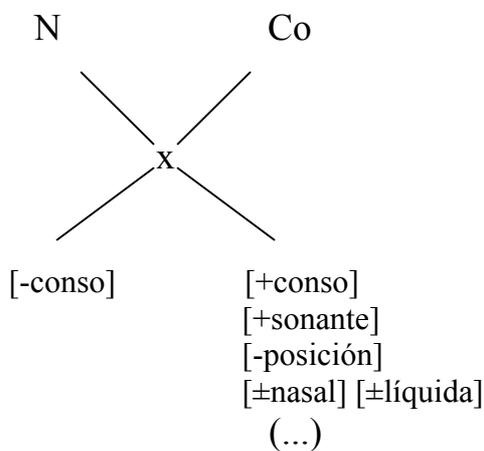
(14)



Nótese que hay solamente un nodo de tiempo. Y que el segmento izquierdo se ha quedado en el nodo raíz pelado, con la especificación de [-consonante] exclusivamente. El resto de rasgos los monopoliza la R, que se adjudica de lleno toda la realización sonora. Ahora bien, el nodo raíz de la vocal tiene que preservarse, porque es ahí donde reside la silabicidad.

Merece la pena ahora que integremos este desglose subsegmental con el esquema general que propusimos para las CS. Tiene los nodos silábicos de costumbre, pero sustituimos los segmentos por sus rasgos constituyentes:

(15)



Aquí no está sólo la consonante R, sino que figura también el nodo vocálico desnudo. Esta configuración nos parece la más fiel para dar cuenta de cómo la CS se asocia al nodo N. No prescindimos enteramente del segmento nuclear, y al mismo tiempo dejamos que la consonante aporte todos sus rasgos de nivel inferior. La R “invade” el terreno del N e impone sus rasgos sobre todo el conglomerado.

Proponemos asimismo que la schwa desprovista de todos sus rasgos de bajo nivel, y limitada a su especificación más fundamental reciba la consideración de **núcleo vacío**. Con ello estamos asumiendo un concepto primordial de la Fonología de la Rección¹⁵, pero introducimos un par de modificaciones sustanciales. En primer lugar, nosotros estamos utilizando rasgos y no elementos como constituyentes del segmento. La geometría de rasgos permite aquilatar en qué consiste exactamente el segmento denominado como núcleo vacío, en tanto que no somos capaces de encontrar un equivalente en la fonología de elementos.

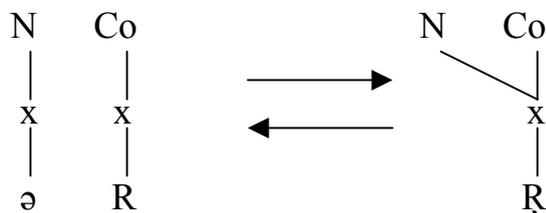
La segunda toma de distancia es que en nuestro marco teórico el núcleo vacío está resultando ser una excepción, mientras que en la FR presenta una distribución muy amplia. Nuestro punto de vista nos mueve a asociar algo de la excepcionalidad de las CS con el ‘núcleo vacío’, precisamente.

La estructura que acabamos de plantear para las CS cuenta, en nuestra opinión, con una serie de ventajas que pasamos a detallar. Uno: Evitamos la inserción directa de una consonante bajo el nodo N, o al menos de una consonante pura. Dos: La presencia del N vacío recoge el carácter **híbrido** de las CS, que son a la vez consonánticas y

vocálicas. En el plano fonológico, es lo que apunta Scheer (2009: 1-2) cuando al hablar de las CS se refiere a su “*hermaphrodite identity: their body is consonantal, but their behaviour is vocalic*”. Y en el plano fonético y en lo que respecta a la composición de los segmentos, recordemos que el estudio de Yu (1999) deja muy clara la naturaleza simultáneamente vocálica y consonántica de la /z/ en *sz* (en tanto que la /s/ del inicio es exclusivamente consonántica, y en *si* la /i/ es puramente vocálica). En definitiva, así es como la CS tiene carácter nuclear y silábico pero no tiene la realización fonética de la vocal.

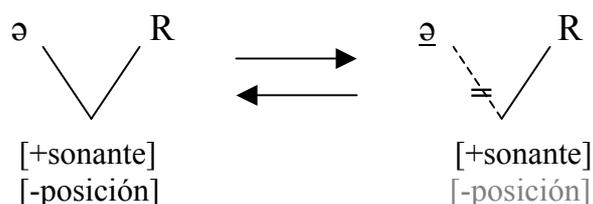
Tres: La Fonología de la Rección (FR) habla de núcleos vacíos, pero se ciñe mucho al plano segmental. Sin embargo, las categorías vacías deben estar lo más fundamentadas que se pueda, y por eso hemos descendido al detalle infrasegmental¹⁶. De este modo somos capaces de explicar cómo puede ser que esté *vacío* (no tiene realización fonética, por falta de rasgos de bajo nivel) un *núcleo* (sí que se conserva el nodo vocálico). Cuatro: La idea de núcleo vacío es un pilar básico para la alternancia entre schwa y cero, de acuerdo con la FR. El nodo vacío favorece la reversibilidad¹⁷ de la regla:

(16)



porque basta con volver a ligar los rasgos para recuperar la əR:

(17)



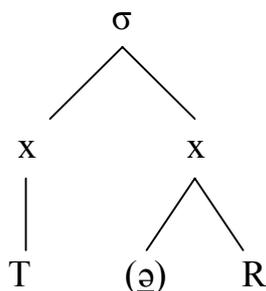
(N. B.: La schwa subrayada sigue la notación de la FR para los núcleos vacíos.) Preferimos nuestra perspectiva, que toma la schwa plenamente expresada como forma de origen, a las que parten de un nodo vacío (o nodo raíz de la vocal) y lo rellenan con reglas de redundancia, de infraespecificación radical o algo similar¹⁸.

En la sección que sigue analizamos un quinto aspecto que nos parece de interés. Nos referimos a la semejanza que presentan las CS con las consonantes africadas. Como veremos a continuación, la estructura ramificada alberga dos segmentos bajo un sola unidad de tiempo.

Las CS como ‘estructuras africadas’

Siempre hemos tenido presente el carácter excepcional de las CS, y hemos planteado que su estructura es un tanto peculiar. En este capítulo hemos ahondado en los detalles y hemos propuesto que en cierto modo la coalescencia de las CS tiene que hacerse sobre un N vacío. En esta sección vamos a plantear que esta estructura excepcional se **legitima** gracias a un mecanismo igualmente excepcional. Para ello tomamos la CS en su conjunto, en una perspectiva más amplia que incluye también al inicio:

(18)



[anotamos el N vacío como segmento subrayado]

Observemos que hay dos ramificaciones; una que pivota en torno a la unidad de tiempo y otra en torno al nodo sílaba. Tenemos la impresión de que el inicio actúa como **legitimador** de la otra estructura, que incluye un N vacío. Si no existiera el inicio, no podría licenciarse la estructura del N vacío.

Es más, nos planteamos que la estructura del diagrama funciona como una africada. El **‘primer tramo’** es la T, que en las africadas es necesariamente una oclusiva, y en las CS abarca un espectro más amplio pero tiene un gradiente de preferencia (donde las oclusivas son las mejores candidatas, y aún más si son insonoras). En las africadas, a diferencia de las oclusivas puras, ese primer tramo ‘duro’ no se relaja de golpe, sino que lo hace a lo largo de un **‘segundo tramo’**, que es una fricativa. Lo que proponemos es que en las CS, la R funciona de manera análoga, como tramo de liberación o relajación de una consonante previa (la C del Inicio).

Esta es la propuesta entonces: las CS vienen a ser como el segundo tramo o **‘resto’** (o apertura de la oclusión) de las consonantes africadas, y para ello necesitan de un primer tramo, que es el inicio y

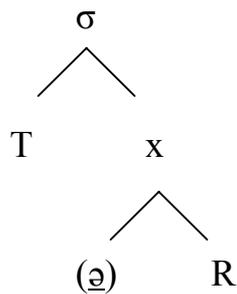
que por eso es imprescindible en estas formaciones. Las CS vienen a ser una suerte de **complemento** de su inicio, que es quien les da paso y las legitima.

La legitimación es de todo punto necesaria, porque africadas y CS incumplen el Principio de Secuenciación Sonora (PSS), con esa segunda parte más sonora que la primera. Por ejemplo, si tomamos *much* segmento a segmento, /m-ʌ-t-ʃ/, el último de ellos es más sonoro que el penúltimo. La secuencia puramente lineal no funciona. ¿Cómo se resuelve este problema? Pensamos que la clave está en reducirlo a un **ámbito local**, y por eso la estructura es tan importante. En las africadas, hay una asociación 2:1, en la que dos consonantes están vinculadas a una sola unidad de tiempo. Así es como se delimita una **cápsula** (*shell*) que a ciertos efectos cuenta como un solo segmento, y que sólo internamente consta de dos segmentos. La coalescencia añade valor a esta manera de computar.

En las CS, la cápsula no está dominada por la unidad de tiempo ‘x’, sino por el nodo sílaba ‘σ’. Es decir, que se ha trasladado un patrón ya existente a otro lugar, de acuerdo con un recurso habitual en la evolución de los seres vivos¹⁹ y que nos parece de lo más aplicable a los mecanismos cerebrales del lenguaje.

En las CS hay que contar con más elementos y ramas que en las africadas, por lo cual la estructura se complica un poco. La asociación 2:1 se forma en dos ocasiones, concretamente. Una es la que une dos segmentos a una unidad de tiempo (igual que en las africadas), y la segunda es la que vincula a la T con la propia CS:

(19)



[Hemos prescindido de algunos elementos (nodos silábicos, unidades de tiempo) para dejar la representación en sus elementos primordiales y seguir mejor el razonamiento.]

La T queda al nivel del aspa de la unidad de tiempo, lo cual apunta a que no tiene que lidiar con dos segmentos a la vez, sino con un solo constituyente. Y además, el núcleo vacío resulta **transparente** y permite que la T tenga acceso a la R y la licencie por contacto directo. Así es como la R puede servir de ‘resto’ de la T.

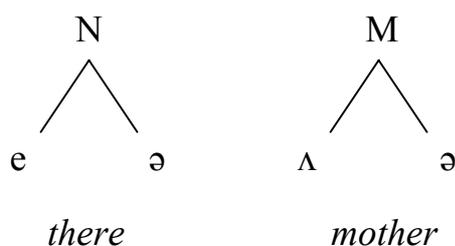
Ahora se aprecia mejor el paralelismo con las africadas, cómo la T licencia al material que le sigue y cómo se forma una cápsula que permite romper con el PSS gracias al ámbito local que delimita. La coalescencia de segmentos (los propios de la CS) debe contribuir a la legitimación, y ello a pesar de no contar con la coalescencia subsiguiente con la T. Pero algo de eso sí que hay en la articulación totalmente homorgánica en las CS más fuertemente silábicas (las del tipo *mountain* /^hmaʊnʔŋ/), en la medida en que la oclusión faríngea se hace conservando la posición alveolar de la lengua.

Un factor que puede contribuir al éxito de las CS, al tiempo que complica su estructura, es el lugar que ocupan dentro del troqueo. En su condición de sílaba débil, la CS actúa como tramo de liberación o

‘resto’ de la sílaba fuerte, y con ello reitera el esquema de las africadas en un nuevo nivel, más alto. Si el troqueo se representa como árbol, se lleva la ramificación hasta el último tramo.

Aún más, la estructura de “africada” propia de las CS deja muy clara su calidad de elemento dependiente, y mucho nos parece que este carácter viene **heredado** de la propia schwa. Al fin y al cabo, la schwa nunca aparece por sí sola, y además ocupa siempre una posición subordinada y **dependiente**. Esto puede suceder dentro de un mismo núcleo N, como ocurre en *there*, o también con vocales de otro N cuando están conectadas por el ámbito de un pie, de una supermora, como ocurre en *mother*. La distinta notación para la sílaba (con árbol) y para los pies (con paréntesis y aspas) pudiera confundirnos, pero realmente es lo mismo: {cabeza + complemento}, donde este depende de aquella:

(20)



Después de recorrer las similitudes, es hora de señalar algunas diferencias importantes entre africadas y CS. En las africadas destacamos su carácter estable y no-marcado, y también el carácter estridente del segundo tramo (incluso en /pf/), que además suele ser sibilante (en /tʃ, dʒ/)²⁰. La estrategia de las africadas para el segundo tramo es un alto nivel de ruido, y la de las CS, la mayor sonoridad. Por

parte de las CS, nos quedamos con su valor nuclear y silábico (que acarrea peso moraico), y lo grande del ámbito en el que se extienden (pivote en el nodo silábico, y alcance hasta el nivel del troqueo). Desde esta perspectiva, las CS son doblemente “silábicas”: por la asociación de la ‘R’ al núcleo, y además, por abarcar todo el ámbito de una sílaba en su estructura completa.

Las africadas no saben nada de la alternancia que tanto caracteriza a las CS inglesas. No nos parece descabellado que a largo plazo sólo sobreviva uno de los términos de esa misma alternancia (el más estable, que sería la pura schwa o el ‘cero’). El otro término se quedaría abocado a la extinción, quizá porque las CS sean estructuras de transición, dentro del marco de todo un proceso evolutivo más amplio.

En definitiva, hay dos claves estructurales que resuelven los grandes inconvenientes de las CS y por tanto las legitiman. Uno, el núcleo vacío, que nos evita el engorro de tener una consonante directamente insertada bajo el nodo N (si conservamos algo de V es mejor), y que es capaz de hacerse cargo de la silabicidad del conjunto. Y dos, la “africación”, que nos evita el engorro de romper el PSS, gracias al carácter local que delimita la cápsula (*shell*).

Insistimos finalmente en la estrecha relación entre el *cómo* y el *dónde*, que esta vez se sustancia en reutilizar componentes ya existentes, pero trasladados a otros niveles. Este mecanismo es un motor fundamental para la evolución de los seres vivos, de acuerdo con la **Genética Evolutiva del Desarrollo** (también conocida como **evo-devo**). En el

penúltimo capítulo de esta tesis volveremos a ver este recurso evo-devo, pero aplicado a las CS atípicas.

¹ Toft comenta que la /n/ silábica es más larga que la /l/ silábica, pero siempre dentro del margen de una unidad de tiempo (*timing unit*); como ella señala, nadie entiende que una unidad de tiempo corresponda siempre exactamente a la misma duración. Pero en líneas generales sí podemos afirmar que las CS son cortas.

² Tanto es así, que Hammond (1999: 202) considera a las sílabas encabezadas por una schwa como *zero moraic*, esto es, nulimoraicas.

³ El único caso donde una schwa nuclear convive con vocales es en los diptongos con [əʊ] (ver Cruttenden 2001: 134-36), pero sospechamos que ahí no hay una verdadera /ə/, porque la vocal de esos diptongos es más tensa que la vocal que conocemos como ‘schwa’. Una conjetura es que la vocal de estos diptongos es la /ɚ/. Pero incluso si se tratara de una especie de ‘schwa tensa’, con el mismo lugar de articulación, evitaríamos la coexistencia de schwa silábica con otra vocal postulando que los diptongos sean subyacentemente /ou/.

⁴ Al fin y al cabo, hay CS sin nodos hermanos. Las CS de coda tienen que remontar su condición de ‘hermanas’ al nodo de la rima (que es la proyección del N, eso sí), y en las CS que ocupan un inicio (en la sílaba siguiente, cuando tenemos CS en contexto vocálico, CS / __ V) ni siquiera se da esta situación.

⁵ Pensemos en vocales poco sonoras en contacto con la schwa: *poor* /pʊər/, *weird* /wɪərd/, *swear* /sweər/ (ejemplos tomados de ‘Centring diphthongs’ en Cruttenden 2001: 141-46). La schwa siempre queda como complemento del diptongo, y la otra vocal es la cabeza, pese a su baja sonoridad.

⁶ El alto grado de afinidad entre la schwa y la /r/ se aprecia en varios espectrogramas de Ladefoged (2001a), concretamente p. 183 (figura 8.14), p. 188 (figura 8.18), y pp. 212-13. Sobre el espectrograma de la figura 8.18, dice Ladefoged que *one of the next most identifiable points is the drop in the third formant below (21) at the end of the word “were”, showing that on this occasion I did pronounce the [ɹ] in this word*. Es decir, que la asimilación rótica se recoge en forma visible en los espectrogramas: la caída del F3 cuando aún se está articulando la vocal (schwa) es una expresión acústica de la coalescencia, en nuestra opinión.

Ladefoged (2001a: 212-13) abunda en esta cuestión: “*Rhotacization is an auditory quality, which, like height and backness, is most appropriately defined in acoustic terms. In a rhotacized vowel (or portion of a vowel) there is a marked lowering of the frequency of the third formant. The frequencies of the first two formants determine the vowel height and backness. The frequency of the third formant conveys comparatively little information about either of these aspects of vowel quality. (...) there is a large fall in the frequency of the third formant in words such as “deer” and “bear”, in which the ends of the vowels are considerably rhotacized in many forms of American English. Furthermore, throughout most of the word “heard”, the third formant may be low, indicating that even at the beginning of the vowel there is a rhotacized quality.*”

La falta de especificidad de la vocal rotificada encaja muy bien con la “neutralidad” de la schwa. Por eso no es casualidad que en las palabras que cita Ladefoged, *deer* y *heard*, la vocal sea larga y tenga un segundo tramo más laxo, que corresponde a una schwa (como en *there*, /^hðeər/). El espectrograma de (2001a: 214, figura 9.12) analiza la secuencia “*the deer heard the bear roar*” y deja ver muy nítidamente cómo es el segundo tramo de las vocales, que corresponde a una schwa, el que tiene la bajada fuerte del tercer formante, que se acopla con la rotización.

⁷ La complejidad en este caso se juega por debajo de los rasgos más generales. Hay una interacción entre los rasgos de posición, por la cual ciertos rasgos de cada segmento formarán parte del nuevo segmento, y otros de esos rasgos no lo harán. La clave está en la geometría de rasgos.

⁸ En concreto, el primer segmento es una oclusiva (bien la /t/ para formar una /tʃ/, o bien /d/ para formar una /dʒ/) que aporta los rasgos de [+consonante] y [+obstruyente], así como [α sonoro] y [coronal]. Y el segundo segmento, la semiconsonante /j/ aporta igualmente [+coronal] pero con el matiz [posterior] (o bien [-anterior]), y además el rasgo [+continuo], que hace que la obstruyente sea una fricativa. Como hemos señalado, el rasgo [estridente] parece ser un rasgo **emergente**, que aparece justamente a raíz de la propia coalescencia.

⁹ La reducción y neutralización se hacen en círculos concéntricos, son centrípetas en inglés. Incluso podemos pensar que el círculo mínimo, cuando se cierra el espacio, es lo que corresponde a ‘cero’.

¹⁰ Gracias al rasgo [tenso] somos capaces de captar el efecto centralizador (que la FR recogía con el elemento @), ya que produce este resultado.

¹¹ La schwa es casi la única vocal que no tiene una alternancia entre forma redondeada y forma extendida, como se aprecia en los diagramas de Ladefoged (2001a). Y podemos derivar la ausencia de [redondeado] a partir de las características generales de las vocales inglesas. En esta lengua, todas las vocales redondeadas son traseras, por lo que cualquier vocal no-trasera debe ser igualmente no-redondeada.

¹² Hay que contemplar la posibilidad de que la /ʊ/ se reduzca directamente a CS, en vista de lo que comenta Ladefoged (2001a: 80): “*A reduced vowel more like [ʊ] may occur the suffix “-ful” as in “dreadful” [ˈdredfʊl], but for many people this is just a syllabic [ɹ] [ˈdredfɹ].*” La cuestión es si forma parte de una reducción progresiva, que pasa por la schwa, o se pudiera ir directamente a la CS.

¹³ En este sentido, Scheer (2009) la considera un tanto aparte, en otro tipo de núcleos, algo así como si no fuera una verdadera vocal. Y asimismo, Ewen & van der Hulst (2000) dicen que *the status of schwa is more problematic. (...) and shows other phonological behaviour which suggests that it forms a set of its own in some respects.* (2000: 18, nota 16).

En otra muestra del carácter problemático de la schwa, estos autores incluyen a la schwa entre las vocales [-tensas] del ‘RP’ (dialecto inglés británico estándar) (2000: 18), pero a renglón seguido le asignan el valor [+RLA] (raíz lingual adelantada, o ATR por sus siglas en inglés) en el dialecto asante de la lengua akan de Ghana (2000: 19-20).

¹⁴ Schane (1995: 597) da por hecho que es la propia schwa la que no tiene partículas, y se limita a un nodo raíz. En concreto, explica que en el sistema vocálico del frisio occidental (una lengua muy cercana al inglés, dicho sea de paso) hay 9 vocales y que la schwa ocupa un lugar muy especial: “*a tenth vowel [ə] (schwa) occurs only short*”, y sobre todo, “*The schwa vowel has a root node with no particles under it*”.

El galés meridional tiene un patrón similar en su sistema vocálico: cinco pares de vocales en oposición larga/corta, y después la schwa por sí sola. (Roca & Johnson 1999: 292-93).

¹⁵ Remitimos al Apéndice I para una explicación sobre los núcleos vacíos en la FR.

¹⁶ Es cierto que Harris (1994: 192) asume un elemento @ latente para los núcleos vacíos. Pero en esos términos cuesta entender que tenga dos posibilidades de expresión fonética, bien como schwa (*fidəl*) o bien como CS (*fidl*).

¹⁷ La reversión del proceso no es universal, por otra parte. Recordemos que hay CS que no alternan en absoluto. Pensemos en *people* o en el sufijo #-ble# (por ejemplo, *possible*), que son muy fuertemente silábicas.

¹⁸ En la **Teoría de la Infraespecificación Radical**, se dice que todo sistema tiene un segmento que está totalmente sin especificar en su forma subyacente. Se trata de una **vocal especial** que se comporta como si estuviera ausente o fuera invisible en procesos tales como la armonía vocálica, y que suele constituir la vocal epentética por defecto (esto es, la vocal que se inserta en los contextos que requieren la presencia de una vocal, pero donde la especificación de rasgos vocálicos concretos carece de importancia). [Véase Ewen & van der Hulst (2000: 77)]. La /i/ en khalkha, la /i/ en árabe marroquí (Kaye), la /u/ en japonés y la /ə/ en inglés y otras lenguas son ejemplos de este tipo de vocal (que no es la misma en todos los sistemas: depende de la especificación que se haga).

Esta teoría nos parece muy interesante, cómo no. Pero no la vemos de aplicación en la lengua inglesa, donde consideramos que la schwa no es epentética, y que es una vocal del plano métrico. Eso sí, la schwa tiene algo de vocal “por defecto”, tanto a efectos sonoros (acústicos) como en su composición de rasgos, pero descartamos que la epéntesis o inserción en la estructura (salvo en casos como *bus* → *buses*, si acaso).

Aunque dado el alto porcentaje de rasgos negativos que caracteriza a la schwa en el sistema vocálico del inglés, podemos reinterpretar la infraespecificación como un ir suprimiendo rasgos de bajo nivel, hasta quedarnos sólo con el nodo raíz, o poco menos.

¹⁹ Recordemos la teoría Evo-Devo que hemos expuesto en el capítulo de Introducción de esta tesis.

²⁰ Un detalle adicional es que los dos tramos de la africada son coherentes respecto al rasgo [sonoro], que viene marcado por el valor que tenga en el primer tramo, [α sonoro]. Sin embargo, en las CS la ‘R’ siempre es [+sonoro], en tanto que el inicio es preferible que sea [-sonoro], aunque cabe también [+sonoro].



7 Cómo se forman las CS (y V)

Hemos llegado al último tramo de la exposición sobre cómo se forman las CS. Desde lo más pequeño ascendemos a lo más grande para ir completando nuestra comprensión del proceso en su totalidad. Este capítulo trata sobre la proyección del material fonológico a cotas más altas y el paso de la representación fonológica a la implementación o realización fonética.

La proyección de las CS

Nuestra hipótesis de trabajo es que el plano melódico contiene material estrictamente fónico, dentro del componente fonológico. Los segmentos y sus rasgos constituyentes son representaciones mentales que provienen de capas más bajas de la corteza cerebral, tanto sensorial como motora (referido a su función acústica y a su función articulatoria, respectivamente). Este material fónico se procesa en el plano funcional (con sus dos subplanos, silábico y métrico) para hacerlo asequible al componente cognitivo del cerebro/mente¹. De esto trata la derivación fonológica, a nuestro entender.

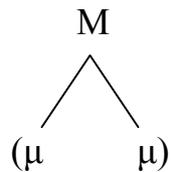
Por la manera en que se organiza la representación fonológica, la arquitectura de planos viene a ser como un **mapa** de la derivación

fonológica. La estructura se construye desde abajo, y se va ascendiendo a niveles cada vez más altos mediante la proyección del material ya existente. Abajo del todo está la **representación subyacente** (en el plano melódico), y a medida que ascendemos en la representación (especialmente, si hay estructura escueta o *Bare Phrase Structure*) se van situando las distintas etapas de la derivación, las sucesivas **representaciones derivadas**. Gracias a los distintos grados de transparencia (recordemos que las proyecciones pueden ser transparentes, semitransparentes y opacas), durante el ascenso hacia niveles más altos se van proyectando ciertas características de los niveles inferiores, y otras se quedan por el camino: esto es lo que se llama **filtrado** (o en inglés, *percolation*).

Pues bien, al llegar al componente métrico es necesario organizar el material silábico en forma de pies. Hemos adoptado la idea del **peso por posición** (Hayes 1989: 356-7), tal que en principio sólo se consignan como moras los componentes de la rima, pero no los del inicio. En el fondo, se trata de una proyección más del núcleo silábico. El material de la cabeza del N es el fundamental, y por eso siempre cuenta como moraico, en todas las lenguas sensibles al peso². El material que le acompaña, sea la rama derecha del núcleo, o sea la coda, puede tener peso métrico o no, según los parámetros de la lengua.

La **proyección** se sigue hasta el nivel más alto. La agrupación de moras produce una ‘**supermora**’, como dijimos:

(1)



y esta supermora es la que encabeza el pie subsiguiente. Como **cabeza** que es, la supermora puede formar un pie por sí sola, o bien tomar un complemento³. En el troqueo, la cabeza queda a la izquierda y el complemento a la derecha. Y en el yambo el orden es el inverso:

(2)



En esta primera aproximación, hemos puesto una mora (sencilla) para el complemento. La supermora consta de dos moras agrupadas, según lo hemos estipulado. Pero ¿de qué está hecha la mora que actúa de complemento? Nuestra hipótesis es que no hace falta especificarlo todo. Basta con que haya material suficiente, y el resto no se computa. En definidas cuentas, estamos ante una proyección semitransparente. Por eso sugerimos que al representar el troqueo (y el yambo, obviamente), la supermora que lo encabeza vaya acompañada de un aspa simplemente, que representa el “material suficiente”:

(3)



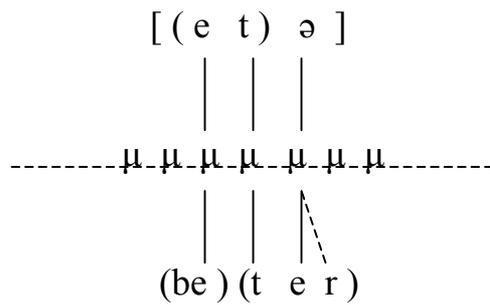
La diferencia en esta representación es que el componente cognitivo no lee más que el “material suficiente” a efectos del peso necesario

para completar el pie trocaico. Ya está. En la posición de complemento, no tiene mayor importancia que la sílaba sea “ligera” (con núcleo y rima sin ramificar) o “pesada” (núcleo y/o rima ramifican; también la coda puede hacerlo). La proyección es semitransparente, y sólo se filtra el material imprescindible para satisfacer una posición de complemento. Sólo se computa una mora. Muy distinto de la cabeza del troqueo, que necesariamente tiene que ser pesada, y que se consigna como supermora (aunque en el momento en que se complete el material suficiente, el resto también se puede dejar como semitransparente)⁴.

Esta forma de proyectar sirve muy bien para la estrategia de **neutralización** que se practica en lenguas como el catalán o el ruso, y también en inglés. La neutralización es una forma de implementar el mecanismo de **luz y sombra** o de **primer y segundo plano** que el componente cognitivo del inglés utiliza para procesar la información fonológica, según planteamos al comienzo de esta tesis. En la parte de ‘luz’ o ‘primer plano’, que es la porción tónica de la palabra o del troqueo, los rasgos de la vocal se filtran todos, de acuerdo con una proyección transparente.

Pero en cambio, los rasgos vocálicos de las posiciones átonas no se filtran enteramente; es un esquema de semitransparencia. En concreto, sólo se filtran los rasgos de nivel más alto, los más cercanos al nodo raíz, y el resto no cuentan para el componente métrico. Así es como se reduce o neutraliza una vocal en su paso al plano superior. Por ejemplo, *better* → *bettər*:

(4)



Hemos representado la schwa y la /r/ unidas a una misma mora de interfaz, ya que en esto consiste la semitransparencia. El sistema sólo lee una unidad de peso, que es lo que necesita para satisfacer la posición de complemento. Por eso la /r/ está ligada al esqueleto superior con una línea discontinua. Queremos hacer notar asimismo la distinta asignación de la /t/ en los dos planos: en el plano silábico está adscrita al inicio de la segunda sílaba, pero en el plano métrico está contribuyendo al peso de la supermora, según lo requerido por el componente métrico. Desarrollaremos esta idea en el capítulo 9, cuando hablemos de la ‘captura de mora’.

Y lo más importante en este momento: el filtrado selectivo de los rasgos de la vocal átona. Así es como la segunda /e/ de *better* se convierte en una schwa cuando pasa al plano superior. Se ha neutralizado gracias a la semitransparencia de la interfaz.

Recapitulemos. El esquema de ‘luz y sombra’ se ha implementado de dos maneras. Por una parte, la aportación métrica de un inicio ha reforzado el peso de la primera sílaba del troqueo, que así puede asumir su carácter tónico. Así resalta más. Y por otra parte, la sílaba átona ha sufrido una pérdida de peso en forma de rasgos no-filtrados.

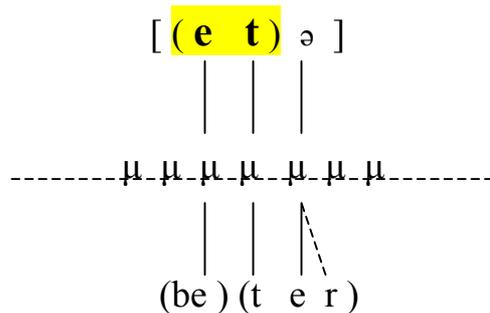
Así queda más en sombra. El doble efecto se aprecia mejor si expresamos el proceso en forma gráfica:

(5)

BETTER → **BETT** əɪ

La porción tónica está aumentada, y al tiempo, la porción átona está disminuida. La combinación de los dos cambios maximiza el efecto de ‘luz y sombra’. Este efecto es lo que habíamos representado en la arquitectura de planos, ni más ni menos:

(6)



Tenemos un troqueo con el patrón [M x], esto es, supermora y complemento. A efectos de computación del peso, el ‘resto’ sólo es sensible a una ‘cantidad suficiente’, que es una mora. En condiciones normales, se tomaría el segmento vocálico como referencia para satisfacer el requisito.

Ahora, la coalescencia de segmentos [schwa + CS] hace que la rima sólo tenga un aspa de tiempo, por lo cual puede filtrarse todo el conjunto. La vocal arrastra consigo a la CS, que se “cuela” al plano métrico. Por eso es relativamente fácil tomarse la licencia de formar

una CS en inglés. Los dos segmentos de la rima, /e/ y /r/ están ligados a la mora del esqueleto superior; el conjunto de los dos queda apto para filtrarse. Y además, el segmento vocálico ha quedado muy mermado (en su carácter vocálico, precisamente) por el filtrado selectivo al que lo han sometido. Una vez metidos en la porción de ‘sombra’, no hay tanta diferencia entre schwa y /r/, y por eso puede licenciarse una CS en esa posición.

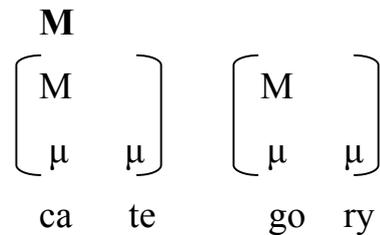
Al reducir a schwa, sólo se dejan pasar los rasgos más fundamentales de la vocal. Esto es lo que se transfiere después a la realización fonética⁵. Y de hecho, una CS es en buena medida esto mismo, porque lo que se transfiere a la realización fonética es la propia consonante, en lugar de la schwa.

Eso sí, es fundamental haber llegado hasta aquí: las CS del inglés se forman en la parte más alta de la estructura. Y también la schwa, por lo cual no sería correcto, en sentido estricto, incluir este tipo de vocales en las representaciones del plano melódico⁶. Esta es la ventaja de hacer una representación integrada de toda la estructura. Por ejemplo, hace falta la coalescencia para que la schwa y la CS se agrupen, que sólo tengan una unidad de tiempo y así poder entrar juntas al nivel métrico. Si nos hubiéramos quedado sólo en este nivel de la derivación, ahora sólo nos quedaría filtrar la CS por pura estipulación.

Terminamos esta sección con un detalle que nos parece muy interesante: vamos a proponer que el mecanismo métrico tiene capacidad **iterativa**. Lo maravilloso de esta capacidad es que le permite operar en otro nivel de la estructura (aunque siempre dentro

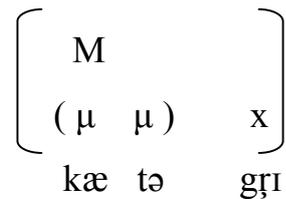
del plano métrico). Vamos a verlo en un ejemplo práctico. En principio, la palabra *category* organiza sus cuatro sílabas en dos troqueos sucesivos:

(7)



Con este esquema, la sílaba ‘go’ porta el **acento secundario**,⁷ y como tal, no es apta para formar una CS. Pero esta estructura puede reducirse y formar una CS, donde /'kætəgɔri/ → /'kætəgɪ/, como la evidencia práctica nos demuestra. ¿Qué esquema métrico se produce entonces? Nuestra intuición es que se traslada el esquema trocaico a un nivel más alto:

(8)



Esta reducción remeda la **conflación de línea** de Halle & Vergnaud (1987)⁸ que asume Hayes (1995: 39, 119) pero en puridad sigue las pautas de nuestro modelo teórico. Porque el primer troqueo ha conservado su estructura original, con las dos moras agrupadas para formar una supermora, igual que antes⁹. La diferencia es que esta vez no forman un troqueo por sí solas (cosa que podrían hacer, como

antes), sino que pasan a formar la parte fuerte de un nuevo troqueo, que ahora tiene una porción débil semitransparente (del tipo ‘x’), en lugar de la mora (del tipo ‘μ’) que había anteriormente¹⁰. Con esto se mantiene la binariedad en una secuencia melódica y silábica de más de dos elementos¹¹. Y sobre todo, a nivel métrico y cognitivo se pone en la línea de ‘sombra’ todo un troqueo completo (merced a la formación de la CS; si no sería imposible), con lo que se facilita el procesamiento de una secuencia tan larga.

Implementación fonética

La representación fonológica está al servicio de la interacción con otros dos componentes del cerebro. Por un lado está la interacción con la función cognitiva, que es la que da sentido a todo el plano superior de la derivación (es decir, el plano silábico y el plano métrico), como hemos apuntado en varias ocasiones a lo largo de esta tesis. Y por otro lado está la interacción con la función fonética, que a su vez consta de una porción motora y otra sensorial. La corteza motora tiene que recibir material de toda la derivación fonológica, y no sólo de la porción más alta de la arquitectura de planos¹².

Hay que integrar todo el material fonológico, de acuerdo con el punto de vista holístico que nosotros mismos venimos sosteniendo. Pero al considerar el proceso en todo su conjunto, las cosas se complican un poco. El componente métrico opera con cierta laxitud en la ‘zona de sombra’ del troqueo, ciertamente, porque sólo hay que satisfacer la condición de ‘cantidad suficiente’. Ahora bien, las derivaciones están obligadas a conservar toda la estructura¹³, de modo que al llegar al

nivel más alto (el métrico) tiene que seguir presente todo el material, hasta sus niveles más bajos. Porque si no, al llegar a la fonética, ¿cómo convertiríamos una /a/ con dos unidades de tiempo en una [a:]?

Traducido a la formación de CS: la lectura métrica semitransparente quizá no aprecie diferencia de peso entre /əR/ y /R/, pero está claro que se pierde cantidad en el paso de /əR/ a /R/. Se pierde un segmento (la schwa), y también una unidad de tiempo. Este aspecto debe tenerse en cuenta cuando llega la hora de llevar la derivación hasta su último extremo, cuando se entra en la realización fonética para transferirlo a la corteza motora. Aquí se transfiere todo el material, y por eso no es igual una schwa que una CS, no es lo mismo [ə(R)] que [R]. Igual que la schwa debe pesar menos que una vocal plena, y esto se obtiene gracias a la integración de todo el armazón estructural.

Otro aspecto importante de la implementación fonética es que la mayoría de las CS del inglés se caracterizan por la **alternancia** entre las dos formas, que son la /əR/ y la /R/; es decir, las dos son posibles. Pues bien, esta oscilación de formas, que hemos venido denominando como ‘**silabicidad débil**’, adquiere nuevos matices y nuevos ámbitos cuando pasamos a la realización fonética. Nuestro análisis nos mueve a sospechar que casi todas las presuntas “CS” tienen en realidad un componente fonético de schwa, aunque sea “microscópica”.

La /əR/ y la /R/ actúan como tipos polares de todo un espectro de realizaciones. Al transferir la /-əR/ al componente fonético, la vocal se articula en toda una gama que va desde una [ə] plenamente pronunciada hasta la forma más levemente esbozada. En un lado del espectro tenemos la realización del núcleo silábico sin reducir,

expresado como vocal plena. Por ejemplo, se escucha con toda nitidez una [e] cuando Robbie Williams canta “*I’m loving angels instead*” o Annie Lennox dice “*there must be an angel (playing with my heart)*”. Y en el otro polo está la elisión de la vocal, bien como CS pura (sin nada de schwa) o como supresión.

Las circunstancias de la elocución y los factores personales condicionan el grado de pronunciación de la schwa (o CS), que en esto tiene algo de aleatorio, impredecible. Pero existe también una influencia de las condiciones de la representación fonológica, que a pesar de ser complejas son más predecibles. En concreto, es más fuerte la incidencia de CS en *cotton* que en *taken*, por ejemplo. Es más, *taken* necesitaría de condiciones favorables (tales como estar ‘**emparedada**’ entre acentos) para poder dar lugar a una CS, mientras que *cotton* lo tiene francamente más fácil. De manera que el término ‘silabicidad débil’ resulta francamente amplio y rico, porque se refiere a todo el espectro cuantitativo de posibilidades de formación de la CS.

Sea como sea, la derivación fonológica se limita a dos tipos discretos o finitos, la schwa y la CS. Pero estas dos formas polares se realizan fonéticamente en todo un continuo sonoro que recorre toda la gama de posibilidades entre los dos extremos, desde la vocal plena hasta ‘cero’. Nos da la impresión, por tanto, de que hay una codificación **de alto nivel cortical** para la fonología y otra **de bajo nivel cortical** para la realización fonética¹⁴. Entendemos que esto mismo es cierto para la **percepción**, y no sólo para la articulación.

Y no es de sorprender, porque esta es la manera de operar de la corteza cerebral en otros ámbitos. En concreto, la corteza visual (que

es la mejor conocida) convierte una gama de percepciones en tipos discretos a medida que se asciende en los estratos de dicha corteza. De acuerdo con el estudio de Rotshtein et al. (2004), la percepción visual de todo un espectro continuo de imágenes (*morphing*), que abarcan todas las mezclas posibles entre el 0% y el 100% de dos imágenes de referencia¹⁵, produce una identificación sólo con uno u otro extremo del espectro en los niveles más altos de la corteza visual. En definitiva, la realización de las CS utiliza un mecanismo ya existente, y por ello es consecuente con la economía del cerebro humano. Las CS son excepcionales en ciertos aspectos (asociar un N silábico a una consonante lo es), pero al fin y al cabo están utilizando los mecanismos ya disponibles, tanto cognitivos como corticales, para su formación.

¹ El plano melódico habla de *quién soy*, siguiendo la retórica de preguntas, y el plano superior (silábico y métrico), de *con quién estoy* (es un plano relacional, todo el componente funcional de la estructura).

² En las notas del capítulo 3 ya dijimos que hay varias muestras de la sensibilidad al peso del inglés. Las recordamos aquí para mayor facilidad de la lectura: Acento sobre sílaba pesada, pies isocrónicos con sílabas reducidas, y monosílabos al menos dimoraicos.

³ Como ya dijimos en el capítulo 3, cuando la supermora forma un pie por sí sola estamos ante un **troqueo simple** y cuando toma un complemento es el **troqueo compuesto**. Podemos añadir ahora, con la perspectiva que nos da nuestro modelo, que el troqueo simple se forma en la línea más básica del plano métrico, en tanto que el troqueo compuesto se forma una línea más arriba.

Una aplicación de este modelo está en los distintos tipos de sufijos, según su comportamiento respecto al acento de palabra. Por ejemplo, un sufijo que reciba acento como #-ese# (como en *Japanese*) corresponde a un troqueo moraico. Un sufijo atractor del acento como #-ic# (*tonic*) o como #-al# (*universal*) está en la parte débil de un troqueo simple (por eso la vocal previa tiende a acortarse, como en *tonic < tone*). Y finalmente, los sufijos neutros como #-ly# (*commonly*) están en el ‘resto’ del pie.

⁴ Este aspecto nos parece de gran relevancia para la teoría: hasta ahora nadie había resuelto el problema de las sílabas “pesadas” a final de palabra en latín (véase el Apéndice III). Se había dicho que no computaban como tales porque eran **extramétricas**, pero no dejaba de ser una mera estipulación.

La idea de semitransparencia o de ‘cantidad suficiente’ para formar el troqueo nos parece que da buena cuenta de por qué no se distingue el número de segmentos de la sílaba final a efectos de computación del peso. El modelo integrado, con la asignación de moras en la parte alta de la estructura, que permite distinguir entre ‘cantidad’ y ‘peso’ es crucial para lograrlo.

⁵ Pero toda la información sigue estando ahí, en los planos inferiores: no se pierde nada durante la derivación. Si fuera necesario, las reglas de redundancia podrían rellenar con lo que haga falta. Pero hay que considerar como ventaja añadida del enfoque integrado este hecho de que toda la información permanece ahí, en la estructura de planos.

⁶ Recordemos que hay CS que se forman postléxicamente, en configuraciones de frase o sintagma, de pies ya formados (*I'll, there'll; get'em*, etc.), por lo cual tienen que formarse necesariamente en niveles altos de la estructura, del armazón fonológico. Es en el nivel postléxico de la derivación, o al menos, cuando ya se han formado los pies.

⁷ En esta representación hay otro detalle de gran trascendencia. La asignación del acento principal es el resultado de un proceso gradual, paso a paso, de abajo arriba. Lo que estamos señalando es que el acento principal de palabra, igual que el acento secundario, son resultado de una estructura de pies previa, que es la que está representada en el diagrama.

Para asignar acentos en *category*, lo primero que tenemos son dos troqueos. Cada uno de ellos proyecta una supermora (M). La asignación del acento principal de palabra (que representamos en negrita) vendrá después. Es la supermora de arriba. Pero sólo está proyectándose desde las supermoras previas:

M
(M M)

Es la iteración de patrones que hemos señalado en el cuerpo del texto. Asignar un acento es como formar un troqueo: sigue el mismo esquema y proyecta un nivel por encima. Pero eso significa que sólo cuentan los elementos que habían proyectado previamente.

Por tanto, no se asigna el acento a una estructura lineal de espas métricas, como es costumbre en la Fonología Métrica, sino que antes deben formarse los pies, y el acento resulta de manera muy natural de esa estructura previa. Creemos que este planteamiento es muy fiel a la idea de las **Columnas Continuas** de Hayes (1995). Al no haber discontinuidad entre niveles, resolvemos con claridad el problema de las sílabas extramétricas, en los términos que habíamos planteado arriba. La sílaba final de *baby* o de *cotton* no son extramétricas en modo alguno, sino que están perfectamente computadas en la formación de un troqueo. Forman parte de él en toda regla. Lo mismo sucede para *category*.

Lo que ocurre es que la asignación del acento principal de palabra se hace con el troqueo ya formado. Esto es, el acento no se asigna a una sílaba, sino a un pie. Esto es fundamental. Con la idea de las columnas continuas en mente, o la simple proyección paso a paso, deja de haber lugar para la extrametricidad.

⁸ Kager explica la confluencia de línea (en Goldsmith 1995: 392), y cita que Halle & Vergnaud la utilizan para sistemas como el inglés, que tienen una regla sensible a la cantidad para asignar el acento principal y una regla insensible a la cantidad para el acento secundario.

También hay borrado de acentos en védico cuando el constituyente que se afija no pertenece al ámbito de las reglas cíclicas (1995: 394-5).

⁹ Sólo a modo de sugerencia: consideramos bien probable que el **Trisyllabic Shortening** (TSS, también llamado *Trisyllabic Laxing*) funcione de esta misma manera. Es el caso de *serene* → *serenity*, *cycle* → *cyclical*, *sane* → *sanity*. Nosotros estamos proponiendo un tratamiento métrico, frente al tratamiento léxico de Kiparsky (1982).

¹⁰ En el paso de *cate-gory* a *category* no hablamos de ‘conflación de línea’, pero sí que se pierde material métrico. En concreto, se pierden dos elementos del segundo troqueo, la supermora y una de las dos moras simples. Al perderse una supermora, la otra que queda sigue proyectándose a un nivel más alto, para el acento de palabra, pero esta vez la proyección es **vacua**.

¹¹ Hay un indicio de que la estructura sigue siendo binaria, pese a contar con tres elementos. La secuencia ‘ca-te-go-ry’, en dos troqueos, permite pronunciar las sílabas de una en una, más lentamente. Lo mismo puede aplicarse a la palabra *Canterbury*, que en *General American* (GA) se pronuncia /¹kænəɾberi/, con dos troqueos (muy isocrónicos) y todas las vocales perfectamente audibles. En cambio, al reducir a CS, nos parece imprescindible cambiar la cadencia. Si decimos /¹kætəgɾi/ o /¹kæntəbɾi/, al menos en algunos dialectos, hay que acelerar la dicción de la segunda sílaba. El esquema rítmico obtenido es más bien [negra + dos corcheas] o bien [(dos corcheas) + corchea] que (tres corcheas) o (tres negras). En definitiva, esquema binario (a base de subdivisiones) y no esquema ternario.

¹² En este sentido, podemos entender que el componente fonético queda arriba del todo en la arquitectura, y que el producto de la derivación fonológica pasa por ahí. Otra posibilidad bien distinta sería la de una salida en paralelo desde cada uno de los planos.

¹³ “*Transformations are structure-preserving*” es una máxima del modelo de **Principios y Parámetros**; por ejemplo, ver Haegeman (1994: 337ss.).

¹⁴ Ramachandran (2011: Cap. 3) dice que hay una ‘división del trabajo’ dentro de la corteza cerebral. Y dice que tiene sentido organizar las regiones en jerarquías, porque a medida que ascendemos en la jerarquía, cada nivel más “alto” realiza tareas más sofisticadas. Pone como ejemplo que la información relativa al color que se procesa en V4 (el cuarto estrato de la corteza visual del cerebro) se transfiere a las áreas más altas del color. Y también la computación numérica (que probablemente sea análoga o tenga una relación estrecha con el lenguaje, al menos con parte de este) parece ser que se realiza en fases o etapas.

¹⁵ Una es Marilyn Monroe y la otra, Margaret Thatcher.

8 Relación inicio/CS: el plano segmental

Los capítulos que siguen van a estar dedicados al papel que juega el contexto en la formación de CS. Dos capítulos van a estar dedicados al contexto izquierdo, y otro más al contexto derecho. Dentro del *cómo*, veremos cuáles son las condiciones que impone el contexto previo, y también veremos nuevos factores cognitivos. Y en lo que atañe al *porqué*, descubriremos un nuevo factor que contribuye a la reducción fonética con la que se forman las CS: la economía articulatoria.

La relación entre las CS y su inicio

El primer dato que manejamos al tratar de la relación entre las CS y su inicio es que en ninguna lengua (germánico, eslavo, bereber) se observan CS en comienzo absoluto de palabra¹. El hecho se puede expresar como prohibición de CS en contexto inicial absoluto:

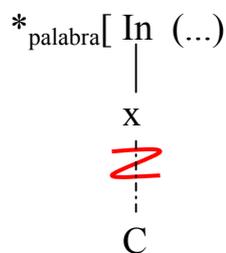
(1)

*CS / [# __] palabra

Aunque se trate de una condición de buena formación o de legitimación, nuestra preferencia es que se formule en términos

positivos, diciendo que las CS iniciales de palabra necesitan un inicio a su izquierda. El requisito se explica en parte por un principio más general de la economía cognitiva. Dado el valor que el Inicio inicial tiene para el reconocimiento de formas léxicas, el sistema se resiste a eliminar material del Inicio inicial de cualquier palabra²; así es como lo entendemos, siempre dentro de la idea de que las CS se forman (al menos en origen) por eliminación de material fonológico (‘reducción’, en sentido amplio). Por lo tanto:

(2)



Esta condición (observable a principio de palabra) de tener un inicio a su izquierda la hacemos extensiva a todas las CS, sea cual sea su posición. Ahora nos basamos en un criterio estructural, que ya habíamos apuntado: la legitimación de las CS mediante una estructura “africada”, donde la CS hace de tramo ‘relajado’ de una ‘T’ previa. No debemos olvidar que para ello la ‘T’ tiene que delimitar un ámbito local en el que incluye a la CS. Porque si no sería difícil explicar que exista una interacción significativa entre inicio y CS

Pero la relación entre inicio y CS parece bastante segura, a la vista de las restricciones que las CS nasales del inglés tienen respecto a la sonoridad y al lugar de articulación de la consonante previa. El dato queda bien patente en Toft (2002), pero además nosotros tenemos nuevos datos que reafirman y amplían la idea de que las CS inglesas

son sensibles al contexto izquierdo. Así lo iremos viendo a lo largo de este capítulo, poniendo de manifiesto una vez más la interacción entre inicio y CS que hemos sugerido³.

Pero, sea como sea, no deja de ser insólito que un inicio afecte al núcleo. No debemos perder de vista que el inicio está considerado como una posición bastante independiente del núcleo, debido a la autonomía fonotáctica que hay entre ambos nodos. Así es como el inicio queda excluido de la rima (donde se agrupan (N-Co)). Pero si los datos sobre CS se salen de la norma (de independencia inicio/rima), habrá que investigar el vínculo que une a las CS con su contexto izquierdo. Vamos a ver los datos con más detalle.

Influencia de la sonoridad

La norma de buena formación silábica dispone que el inicio sea menos sonoro que el núcleo. Las CS no son una excepción en este sentido: la **distancia sonora** entre inicio y CS es favorable o incluso necesaria para la formación de esta última⁴. Este argumento se basa en el amplio muestreo que realizamos en el diccionario *Everyman's*, y que nos mostró que las CS nasales tienen clara preferencia por los inicios obstruyentes a su izquierda (si bien las líquidas apenas tienen restricciones)⁵. Poniendo los datos observados por Toft en relación con la Hipótesis TVR⁶ (que desde la teoría apunta a un contraste In/Co), la distancia sonora entre inicio y CS es insoslayable. La noción general que derivamos es que cuanto menos sonoro sea un inicio, más favorable será para la CS, sea esta del tipo que sea.

La distancia inicio/CS se puede cuantificar, como ya hicimos con las consonantes candidatas a CS, utilizando la fórmula $\Delta(\text{SON})$. Sólo que esta vez $\Delta(\text{SON})$ se refiere a la distancia sonora del contexto izquierdo, la que guarda el inicio con la CS:

(3)

$$\Delta(\text{SON}) = (\text{SON}_{\text{CS}} - \text{SON}_{\text{Inicio}})$$

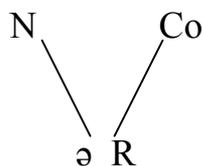
La escala numérica nos permite establecer un requisito de mínima distancia sonora entre la CS y su inicio previo. De entrada, asumimos que $\Delta(\text{SON}) > 0$, que la CS tiene que ser más resonante que su contexto izquierdo; aunque ya veremos que hay excepciones, ocasiones donde se aceptan $\Delta(\text{SON})$ negativos. Y aún más, la fórmula $\Delta(\text{SON})$ lleva a una relación de implicación lógica, donde Inicios oclusivos \supset fricativos \supset sonantes. Queda delineada entonces una gama de inicios para las CS, donde las oclusivas son el contexto más favorable, seguidas por las fricativas y después por las sonantes⁷.

Otra aplicación de la medida $\Delta(\text{SON})$ es formular la mayor sensibilidad de las CS nasales respecto a las líquidas como $\Delta(\text{SON}_{\text{NAS}}) \geq \Delta(\text{SON}_{\text{LÍQ}})$.⁸ Y también gracias a la distancia $\Delta(\text{SON})$ podemos ver la relación que la CS guarda con los distintos nodos. Su afán de lejanía con el inicio contrasta y al tiempo se complementa con la afinidad sonora de la CS por su N, que en el capítulo 1 habíamos formulado mediante una $\Delta(\text{SON})_{\text{RIMA}} = (\text{SON}_{\text{V}} - \text{SON}_{\text{CS}})$.

Vamos a volver sobre algunos conceptos fundamentales. Recordemos cómo la afinidad sonora es la que permite la coalescencia entre la CS y el núcleo. Las dos partes contribuyen a este efecto, porque la

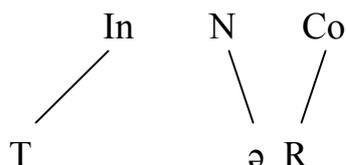
consonante candidata a CS tiende a ser lo más resonante posible (por lo cual tienen franca preferencia las sonantes). Y al tiempo, el N al que se asocia la CS viene de estar ocupado por una ‘vocal fría’, que es la de menor sonoridad entre todas las vocales. El resultado es la combinación perfecta de (ə-R), que contiene dos segmentos que son vecinos en la escala de sonoridad. Ya propusimos que esa afinidad sonora CS/N se representase gráficamente mediante ángulos de enlace:

(4)



Los ángulos de enlace son expresión de la afinidad o cercanía entre los componentes de niveles más bajos, y pueden dar una idea muy gráfica de la propia estructura silábica, o incluso de los procesos de asimilación o disimilación. Con estos mismos ángulos de enlace podemos representar gráficamente todo el conjunto, para incluir la afinidad sonora (N-CS) y la distancia que separa al N de su inicio precedente:

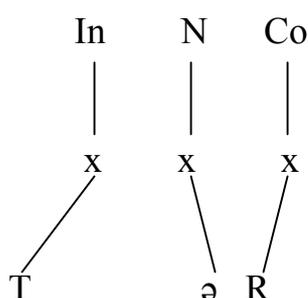
(5)



Este diagrama está simplificado para recoger sólo el concepto fundamental, pero también podemos dar una imagen más precisa: al

incluir las unidades de tiempo (*timing units*) se aprecia mucho mejor cómo los nodos silábicos (In, N, Co) están siempre equidistantes entre sí, y lo que varía son solamente los ángulos de enlace, es decir, la distancia entre los segmentos del plano inferior (y por ende, de sus elementos o rasgos constitutivos):

(6)



Visto así, los ángulos de enlace están reflejando del grado de **afinidad** que puedan tener los rasgos de dos segmentos fonológicos **adyacentes**. Para que una CS sea compatible con su N, debe tener una afinidad de rasgos, tanto en lo referido al lugar de articulación como al grado de sonoridad de un segmento⁹.

Un par de puntualizaciones antes de terminar esta sección. La primera es que no se puede formular todo lo relativo al contexto izquierdo en términos de $\Delta(\text{SON})$; esto es evidente¹⁰. Y sobre todo, un detalle muy interesante para el desarrollo teórico: la propia regla de $\Delta(\text{SON}) > 0$ (que dicta que la CS sea más resonante que su contexto izquierdo) tiene excepciones. En concreto, el inicio /r-/ produce silabicidad débil [r(ə)n], según el *Everyman's*, en vez de la no-silabicidad que cabría esperar de una $(\text{SON}_{\text{CS}} - \text{SON}_{\text{Inicio}}) < 0$. Es curioso que la /l-, siendo menos resonante que la /r-, sí que respete el criterio de diferencia sonora y sólo produzca formas no-silábicas. Pero lo fundamental para

la teoría es que cuando tenemos un inicio /r-/, la sílaba que se forma con la CS es (r̩), con la paradoja de que el inicio es más resonante que el núcleo (cuando este debería ser la porción más resonante de una sílaba).

También se observan CS débilmente silábicas después de un diptongo, que es aún más resonante que la líquida /r/. Eso sí, el diptongo está en el núcleo de la sílaba previa y no en el inicio de la CS. Aun así es otro ejemplo de contexto izquierdo más resonante que la CS. Es posible que sea más importante el apoyo que el diptongo da a la CS en su calidad de contexto izquierdo, pero no es fácil de explicar. Volveremos sobre esta cuestión en el capítulo que viene, por lo que no vamos a profundizar más en ella¹¹, y de momento volvemos nuestra atención hacia el otro factor del inicio que afecta a las CS, que es el lugar de articulación.

Influencia del lugar de articulación

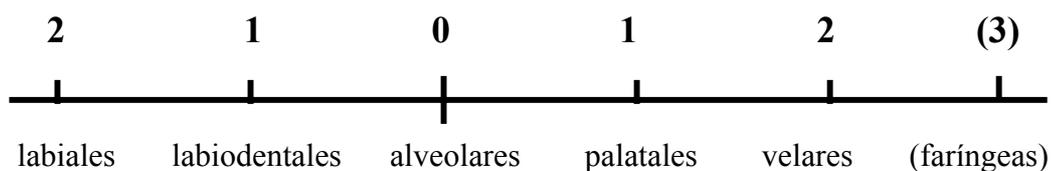
La sonoridad no es el único factor de influencia del inicio sobre la CS. Como apunta Toft, también el [LUGAR] donde se articula el inicio previo afecta a las CS, en especial las CS nasales. Y todo lo que era distancia o contraste inicio/CS al hablar de la sonoridad pasa a ser cercanía o afinidad cuando hablamos del **lugar de articulación**. En concreto, la nasal alveolar /n/ tiene mucha más probabilidad de ser silábica cuando va precedida de un inicio **homorgánico**. La /n/ es fuertemente silábica cuando su inicio es [+coronal]; pero cuando es [-coronal], la silabicidad se resiente y sólo alcanza a ser débilmente silábica o no-silábica (Toft 2002: 117-18). Aún podemos afinar un

poco más y decir que el contexto más favorable para /n/ es el de las coronales anteriores, de acuerdo con el *Webster*. Y por lo que respecta a la otra CS nasal se aplica la misma regla: la labial /m/ prefiere un inicio [labial].

Sin embargo, también el lugar de articulación parece estar sujeto a rupturas de la norma. En concreto, los inicios con /v-/ y con /f-/ tienden a formar CS fuertemente silábicas (*proven, haven't, seven(teen), oven, govern(ment); orphan, hyphen, o siphon*, entre otros). Claro que también hay nasales débilmente silábicas con /v-/ y /f-/, pero los ejemplos de arriba muestran una tendencia hacia la silabicidad fuerte. Al tratarse de inicios no-coronales, parece que se rompe con la regla que habíamos enunciado en el párrafo anterior. ¿Es que no se ajusta al esquema de las clases naturales?

Nos parece muy valioso el concepto de **coronalidad**, y por eso queremos mantenerlo, pero sin quedarnos en la mera distinción [\pm coronal]. La alternativa que proponemos es formular el rasgo [LUGAR] en una escala cuantitativa: en el centro están las coronales anteriores (que son justamente el contexto más favorable), y desde ahí nos movemos indistintamente hacia adelante o hacia atrás:

(7)

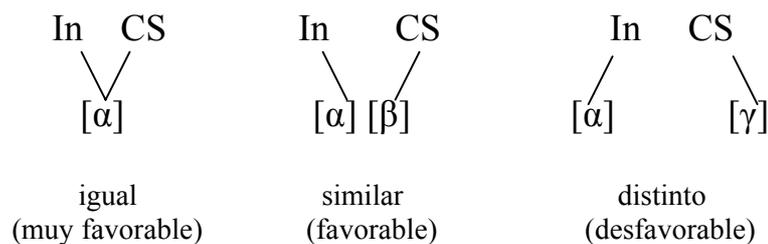


Dicho de otro modo, medimos exclusivamente la distancia con respecto a las alveolares. Así, las palatales (pese a ser coronales) están a la misma distancia de las alveolares que las labiodentales (que no son coronales, empero). Y las labiales están a la misma distancia que las velares de las labiodentales.

Esta escala cuantitativa hace predicciones más precisas respecto a cómo las clases naturales condicionan el grado de silabicidad de las CS. Así, el contexto izquierdo más favorable para la /ŋ/ corresponde a los inicios coronales anteriores, que están en el centro de la escala. A un paso del centro quedan tanto las coronales posteriores como las labiodentales¹², y por ello son algo menos favorables. Siguiendo el camino, labiales y velares quedan ya a dos pasos, y son francamente desfavorables. Finalmente, la faríngea /h/ queda a tres pasos y es el contexto más dificultoso, porque veda el paso incluso a la irrestricta /ʃ/.

Otra ventaja de utilizar una escala cuantitativa es que al expresar de manera precisa la proximidad entre dos segmentos, puede trasladarse también a los ángulos de enlace, como ya habíamos hecho para la sonoridad:

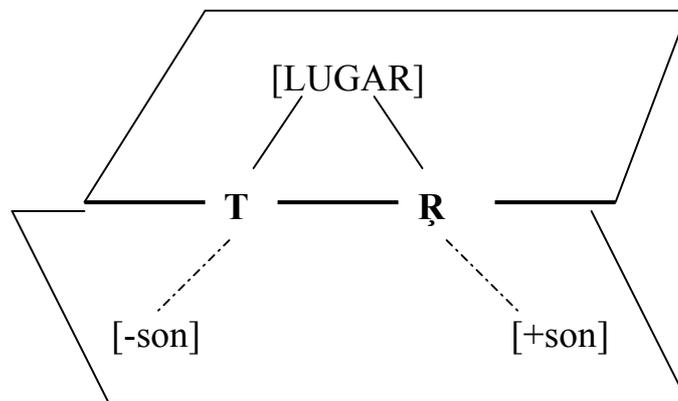
(8)



La escala cuantitativa que acabamos de proponer se basa en el sistema de las **clases naturales**¹³. Si obviamos el pequeño escollo de la coronalidad que hemos comentado, los rasgos de [LUGAR] tienen un buen correlato con la escala, y por ello nos parecen muy adecuados para tratar el comportamiento de los inicios respecto a las CS.

Otra ventaja de la representación con rasgos es que se aprecia muy bien la disociación entre cercanía (para el lugar) y distancia (para la sonoridad) que caracteriza al contexto izquierdo. Porque en la representación autosegmental, los rasgos constitutivos de los segmentos se distribuyen en planos independientes. Al hacer una representación gráfica, distribuimos cada aspecto en su propio plano infrasegmental, y así vemos que no existe contradicción por las distintas características de lugar (cercanía) y sonoridad (lejanía)¹⁴:

(9)

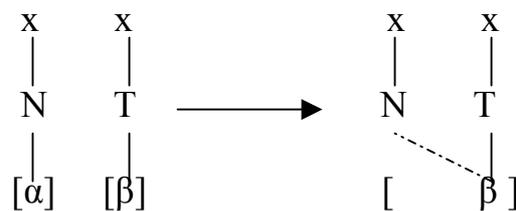


Asimilación Inicio/CS: las antigeminoides

Acabamos de ver cómo un inicio homorgánico es lo más favorable a la hora de formar una CS nasal. Pues bien, esta coincidencia máxima puede obtenerse por **asimilación** del lugar de articulación de la CS con el de su inicio. En concreto, la /ŋ/ asimila el punto de articulación de la obstruyente ‘T’ previa, de modo que [ŋ] → [m̩] en *carpenter*, *happen*, *perpendicular* o *davenport*, según el *Everyman’s*¹⁵. Lo mismo sucede con [ŋ] → [ŋ̩] en *bacon*, y así lo asume también Cruttenden (2001: 293) en la pronunciación relajada de *I can try and book some seats round the corner*.

Esta identidad entre CS e inicio no es subyacente sino derivada, ya que es producto de un proceso de asimilación. Por su dirección es **progresiva**, y sólo en esto se diferencia de la asimilación de los cúmulos N/T (nasal/obstruyente), donde la /n/ subyacente asimila **regresivamente** el lugar de articulación de la obstruyente:

(10)



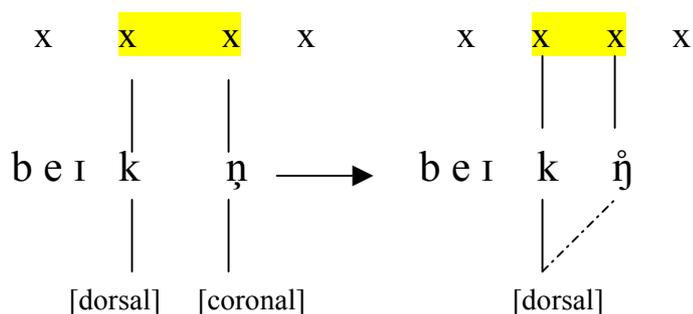
T = obstruyente, N = nasal; α y β son lugares de articulación

Harris se refiere a los cúmulos homorgánicos (N + T) con el nombre de **geminadas parciales** (1994: 69, 76-80). Pero dado que las CS que estamos describiendo son su contrapartida (T + N), nos gustaría que la nomenclatura expresara esta relación. Por eso proponemos reducir el

término de Harris a una sola palabra, dejándolo en ‘**semigeminadas**’ o ‘**geminoides**’. Y así las CS que estamos describiendo, en tanto que representación simétrica de ellas, se pueden llamar sencillamente ‘**antigeminoides**’.¹⁶

También debe de haber una simetría especular en lo que respecta a los nodos silábicos. Las geminoides ocurren en una secuencia (coda + inicio), y las antigeminoides parten de un (inicio + coda) en principio, aunque después tiene que abarcar también el núcleo, es decir (inicio + núcleo-coda). Esto implica que en las antigeminoides existe una fuerte relación (inicio/rima), y como tal se vuelve a romper con el esquema de autonomía entre inicio y rima que es propio de toda sílaba. Nuestra interpretación es que la asimilación desde el inicio hacia la CS tiene que hacerse tautosilábicamente, y casi con certeza, en condiciones de contigüidad estricta. Es necesario que al menos los nodos esqueléticos o unidades de tiempo sean adyacentes; por ejemplo, en la asimilación de *bacon* [ˈbeɪkŋ] → [ˈbeɪkŋ] :

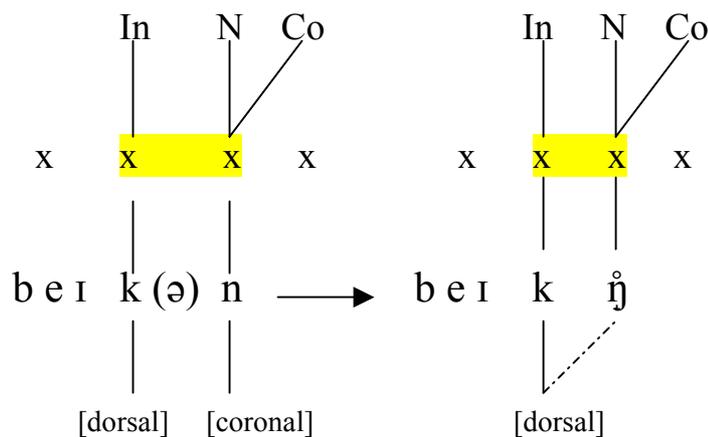
(11)



También es más que probable que tengan que estar adyacentes los rótulos de nodo, es decir, que la CS esté ya en el nodo N cuando se produce la asimilación. Por lo tanto, la asimilación ocurre después de la formación de la CS, dado un sistema de reglas ordenadas¹⁷. Y en

consecuencia, las antigeminoides tienen la peculiaridad de que la nasal está asociada a un N y no simplemente a la coda (que no deja de ser su nodo de origen). Por ello se distinguen un poco más de las geminoides, y más todavía si contamos con que al ser silábica la nasal, tiene peso propio:

(12)



En este nuevo diagrama hemos incluido una schwa en la forma sin asimilar, /^hberk(ə)n/, mientras que en la forma ya asimilada hemos prescindido de cualquier schwa, ni siquiera mínimamente esbozada. Porque otro rasgo distintivo de las antigeminoides es que no admiten el menor grado de schwa. ¡Nada en absoluto! El dato es además muy coherente con la condición de adyacencia estricta que habíamos estipulado.

El fenómeno debe entenderse descendiendo al nivel infrasegmental y a la realización fonética. Es ahí donde observamos que en las antigeminoides la obstruyente va seguida por un ataque glótico [ʔ], que es el que sirve de **ataque** a la nasal. (Obsérvese que aquí el ataque no es el Inicio: nos estamos refiriendo ahora a lo que en inglés se

llama *ictus*.) Este mismo esquema lo veremos en el próximo capítulo con *mountain* [ˈmaʊnt · ʔn̩], con la única diferencia de que *mountain* sólo es posible con oclusivas coronales (y la nasal coronal, claro).

El ataque glótico [ʔ] es necesario para marcar la nasal respecto a la oclusiva, porque si no sería imperceptible. Lo que se hace es emitir la nasal desde el ataque glótico, en un único **gesto articulatorio**, ya que la obstrucción propia de la nasal se hace aprovechando la que ya se había articulado para la obstruyente. Así que en ambos casos, tanto el de *mountain* como el de *carpenter*, lo que tenemos es un ejercicio de **economía articulatoria**, ya en el componente fonético de la derivación¹⁸.

Y naturalmente, esa economía se la pueden permitir hablantes y oyentes porque la **reconstrucción** de la forma original es relativamente fácil. Desde el punto de vista fonológico, sólo hemos dado un pequeño paso desde la schwa hasta la CS; un solo paso y muy pequeño, que hace que las dos formas, subyacente y derivada, sean muy similares. También se sigue el criterio de aplicarlo a palabras de uso frecuente y de un cierto tamaño, de manera que no haya equívocos. Y por supuesto, el proceso es opcional, lo cual puede formularse también como reversibilidad o como alternancia de formas. Así se “conserva” siempre la forma original con schwa. Porque estas CS, a pesar de ser totalmente silábicas, tienen todas una contrapartida donde la schwa está expresada con toda su plenitud. Por ejemplo,

(13)

ˈmaʊntʔn̩ ⇌ ˈmaʊntən

Hay una interacción de fonética y fonología en la formación de las antigeminoides. La motivación está en el componente fonético, según hemos sugerido, pero al tiempo se induce un cambio en la derivación. Tiene que traducirse el gesto articulatorio al esquema de derivación fonológica integrada que hemos propuesto, abarcando todos los planos. Ya hemos planteado la formación de CS como derivación semitransparente, con aplicación de filtros en el ascenso de rasgos (o elementos) hacia el plano métrico. Pero esos filtros son opcionales, de modo que siempre se puede pasar a una derivación transparente, con la schwa. O sea que para pasar de la CS a la schwa basta con suprimir la acción de los filtros, y así volver a la transparencia, en función de lo que decida el componente cognitivo. Bastante sencillo, en realidad. Pero hace falta utilizar todo este esquema holístico del armazón fonológico, porque de lo contrario nos veríamos obligados a estipular mucho más.

En las antigeminoides se va de la fonética a la fonología, por tanto, pero también en sentido contrario. Porque si la fonética es capaz de inducir un cambio en la fonología, lo hace gracias a que el cambio ya se ofrece antes desde la fonología, basándose en las posibilidades que existen. Si no hubiese afinidad de rasgos o elementos, y proximidad estructural absoluta, la opción nunca estaría disponible, y sería absurdo que la fonética viniera a imponer nada. Así que es necesario prestar atención a las características de toda la derivación fonológica¹⁹. Para formar antigeminoides empezamos por la base de la pirámide, por la afinidad de rasgos o elementos infrasegmentales. Desde tan abajo. Y luego tenemos que ir bastante más arriba para inducir una aplicación de filtros de semitransparencia entre el plano

silábico y el métrico, o dentro del propio plano métrico. Así de sencillo, pero así de complejo también.

Al mismo tiempo, es la fonética la responsable de un fenómeno fonológico. Tenemos que empezar por la articulación de dos segmentos concatenados, ‘T’ y /^hn/, y luego dirigirnos a la porción infrasegmental de la representación fonológica, porque es justo ahí donde se desarrolla este proceso. Si nos fuéramos a otro nivel, por ejemplo el de los segmentos, ya no explicaríamos por qué estas nasales son tan fuertemente silábicas; tendríamos que limitarnos a estipular sin más que eso es así, de manera arbitraria. Pero la explicación tiene que ver con el lugar al que remitimos un hecho. Y nuestro modelo integrado de toda la representación fonológica hace hincapié justamente en *dónde* sucede cada cosa. Entendemos que nadie se lo ha planteado exactamente en estos términos, y pensamos que por ello puede ser una aportación interesante, más allá del hecho concreto de las CS.

Las líquidas y el contexto izquierdo

La mejor aptitud de la /l/ para ser silábica (por su mayor sonoridad, respecto a la /n/) se refleja, como es de esperar, en su menor sensibilidad a la influencia del contexto izquierdo, en su mayor tolerancia. No está sujeta a restricciones cuando el inicio es menos resonante, y aún alcanza a ser débilmente silábica con inicios más resonantes. Este segundo caso implica una $(SON_{CS} - SON_{Inicio}) < 0$, un inicio más resonante que su núcleo. Parece que va contra la norma, pero es lo que hay: vemos / $(\emptyset)l$ / en contexto /r__ (*barrel, carol*,

herald, epidural), y aún más sorprendentemente, con un contexto vocálico como /w__ (*narwhal, Cornwall* o *Maxwell* son -(ə)l, si bien *Cartwell* es [-wel]). Queremos hacer notar que la /{r,w}__ puede estar igualmente en la rama derecha de un inicio ramificado, como en *bilingual, sequel; neutral, sepulchral, April* (según el diccionario *Everyman's*). Y aún más, puede no estar ni siquiera en el inicio, porque en el contexto [/diptongo__] hay que asumir que el diptongo ocupa el N previo, y que el inicio de la CS está vacío o ni siquiera existe. Por consiguiente, lo que está en juego no son los nodos o la estructura silábica, sino la adyacencia.

En lo respectivo al lugar de articulación, la /l/ es menos sensible que la /n/, pero no es del todo invulnerable. Toft (2002: 116ss.) muestra cómo la /l/ mantiene su silabicidad con un buen rango de inicios a su izquierda, pero nosotros no hemos encontrado ninguna l/ŋ __ en inglés, que en cambio sí que existe en alemán dialectal, p. ej. *angeln* ['aŋlŋ] 'pescar'. En cambio, para ser contexto de la /l/ inglesa, la /ŋ/ siempre forma un complejo con la /g/, como vemos en *single* o *angle*.

Tampoco hemos encontrado ningún caso de l/ h__ en nuestro estudio, aunque en checo sí que es posible, por ejemplo en *mohl* ['mɔhl] 'podría' (Scheer, 2003b: 3). La ausencia de l/ h__ en inglés encaja muy bien con la escala de distancia que habíamos propuesto: la /h/ no sirve como contexto izquierdo porque está a tres pasos de distancia del centro ('coronal anterior'), que resulta ser demasiado lejos para el inglés (aunque no así para el checo). Queda claro entonces que las líquidas toleran más distancia en la escala de inicios, pero su inmunidad no es completa.

La escala cuantitativa mide la distinta sensibilidad de CS líquidas y nasales en función de los pasos de distancia que son necesarios para perder la silabicidad. También sirve para medir la diferencia entre inglés y checo, que intuimos que pudiera estar fijada por un parámetro, pero además tiene que derivarse de las características de los segmentos, igual que la diferencia entre líquidas y nasales corresponde a su distinta sonoridad o sus componentes infrasegmentales.

Hablando de líquidas y nasales, se da la paradoja de que estas son las únicas que producen la versión más fuertemente silábica de todas (esa que no admite el menor grado de schwa), siendo como son peores candidatas a CS. Porque la /l/ no forma antigeminoides, aunque sólo sea porque no puede adaptarse nunca al punto de articulación vecino; recordemos que sus alófonos ingleses ‘claro’ y ‘oscuro’ dependen únicamente de su posición en la sílaba (inicio y coda, respectivamente).

Sin embargo, nos ha parecido que la /l/ tiende a ser fuertemente silábica cuando está precedida por un inicio (bi)labial. Planteando la alternancia schwa/cero en forma de porcentajes, la variante más silábica debe de llevarse una proporción muy alta. Por ejemplo, en *people* está muy escorada hacia la silabicidad, y muy difícilmente admite una schwa. Lo mismo podemos decir del sufijo {-ble}, que no se presta mucho a la alternancia schwa/cero. Pensamos que en *people* o *-ble* hay cierta identidad de lugar entre inicio y CS, y que así se forman cúmulos bastante homorgánicos. A pesar de que la /l/ no es *per se* labial, la /l/ de coda es muy ‘oscura’ y muta con facilidad a /ʊ/

(como hace en *Estuary English*), por lo cual debe de tener influencia el rasgo [labial] (o bien, del elemento U en la FR).

Desde el punto de vista articulatorio, en los cúmulos *pl-* o *bl-* se repite el esquema de articulación “**africada**”: primero se articula la oclusiva, y según se relaja la oclusión hay una salida lateral (la /l/), todo ello en un solo chorro de aire. Es otro gesto de **economía articulatoria**, no muy distinto del que justifica la existencia de las antigeminoides. En definitiva, hay patrones que se repiten a lo largo y ancho del espectro de las consonantes silábicas del inglés.

La conclusión es que la CS líquida /l/ es francamente versátil respecto a su contexto izquierdo. Su silabicidad se ve favorecida en los cúmulos /pł/ o /bł/, que producen una versión análoga a las antigeminoides. Y por la parte negativa, tiene sólo dos segmentos que impiden su silabicidad, que son /h/ y /ŋ/.²⁰

La ‘distancia cero’

Terminamos este capítulo con un efecto peculiar del contexto izquierdo. Partimos de la hipótesis TVR, que suscita un **contraste** entre T/R, o entre In/Co. Colegimos que a menor contraste, peor sílaba, y de ahí que la peor de todas deba ser la que tiene el inicio y la CS completamente idénticos, o sea, $C_{\text{Inicio}} \equiv CS$. Vamos a llamar ‘**distancia cero**’ a este efecto, que en realidad no es más que un aspecto del **Principio del Contorno Obligatorio (PCO)**²¹. El resultado es que ninguna candidata a CS, sea cual sea, puede ocupar jamás el núcleo silábico si tiene a su izquierda una consonante

idéntica a ella. Lo vemos igual para la /r/ (*emperor, sorcerer, explorer* y *clearer*, o *there're*)²², para la /l/ (que no puede ser silábica en *parallel* o *Hillel*) y para la /n/ (incluso en alemán, que es donde la /n/ goza de más libertad para ser silábica, pero nunca lo es en *weinen, können*, o *Tannenbaum, Birnen*). Debe quedar claro que la ‘distancia cero’ no alude simplemente a una $\Delta(\text{SON}) = 0$, sino a la identidad total y absoluta entre inicio y CS: coinciden todos los rasgos sin excepción²³.

Hay también una versión más “débil” de la distancia cero, donde inicio y CS difieren en un solo rasgo. En concreto, el contexto / m ___ suele bloquear la silabicidad de la /n/, y de ahí que no se diga *I amn't*; pero no siempre es inadmisibile [ŋ] /m___, p. ej. en *rhythm'n'blues*. Y el bloqueo tiene un matiz cuantitativo, porque se atenúa cuando hay condiciones facilitadoras. En concreto, *human* es claramente no-silábica, pero si se incrusta en una frase como *Human Rights Watch annual conference*, que la diluye entre otros acentos más fuertes, entonces puede insinuarse una CS. En cuanto a la líquida /l/, también aquí lo tiene más fácil que las nasales, y no tiene problema en ser silábica detrás de una /r/ (p. ej. *coral*), por lo cual inferimos que el efecto no tiene nada que ver con las clases naturales.

El fundamento de la ‘distancia cero’ quizá se encuentre en restricciones de nivel bajo (esto es, más próximo a la materialización fonética) y probablemente sea análogo al que obliga a insertar una vocal epentética en los plurales, genitivos y tercera persona del singular cuando la base termina en sibilante. Se diría que el impedimento es fonético (acústico y/o articulatorio). Por lo tanto, nos parece suficiente con establecer un filtro de interpretabilidad, y

dejamos de lado cualquier restricción estructural, de rección, o de algún rasgo abstracto. Pero aún se debe recoger ese impedimento fonético en la fonología, de alguna manera. Entendemos que a tal fin basta con prohibir la adyacencia de dos consonantes iguales en la misma sílaba²⁴, siguiendo las pautas del PCO, como señalamos al comienzo de esta sección.

Un detalle crucial es el **ámbito** donde se aplica la restricción; una vez más, el *dónde* es fundamental. Porque si las consonantes adyacentes estuvieran en sílabas contiguas, tendríamos simplemente un par de geminadas, que son perfectamente legítimas. Esta es razón de que la ‘distancia cero’ no tenga efecto alguno sobre las **falsas geminadas**, que se forman en el contexto derecho a ambos lados de una linde de palabra. Es elocuente la comparación entre *parallel* y *bottlelike*, o también entre *exponent* y *evenness*, *openness* o incluso *cottonmouth*, *enlightenment*²⁵.

Al centrarnos en los ámbitos de actuación de las reglas, apreciamos una contradicción: habíamos visto que en las líquidas con inicio más resonante no tenían importancia los nodos silábicos, sino única y estrictamente la relación de adyacencia. Pero ahora, en la ‘distancia cero’ sí que cuentan los rótulos que delimitan el ámbito de la sílaba. De momento sólo somos capaces de constatar la diferencia de criterio que rige los dos fenómenos, y confiamos en que la labor investigadora acabe por dar una solución satisfactoria a esta aparente incoherencia.

¹ En croata encontramos *rt* (cabo, en el mar) [tomado de Scheer (2007a)], así como *rz* (honor) y *rvat* (argumentar, discutir) [tomados de Blaho (2004)]. Pero mucho nos tememos que la ortografía engaña y que no son verdaderas CS, sino que tienen una schwa previa: es lo que hemos escuchado nosotros en hablantes nativos de primera mano.

² Por ejemplo, se puede eliminar el inicio intervocálico en *cantado* → *cantao* o en *candado* → *candao*, pero no el inicial de palabra, porque nunca se lleva a cabo *candado* → **andado*, ni *cantado* → **antado*. Es probable que haya un motivo común con el principio de maximizar los inicios.

³ Esta dependencia entre ataque y núcleo pone en entredicho el punto de vista de la Fonología de la Rección contrario a cualquier dependencia fonotáctica entre el núcleo y su inicio (que es a su vez lo que justifica la secuencia “In-N-In-N-In-N-(...)” que encontramos en los diagramas de esta corriente teórica).

⁴ Esta suerte de disimilación quizá se deba a la baja sonoridad de la CS, que requiere entonces que su inicio sea aún menos resonante para dejar clara su condición de tal. Con ello queda rebajada la sonoridad de toda la sílaba en la que se forma la CS. Esto es, $\Sigma(\text{SON})_{\text{cs}} < \Sigma(\text{SON})_i$.

⁵ La FR pondría objeciones a un inicio más resonante, por las condiciones de rección. Pero desde nuestro punto de vista, el N sigue estando allí (por lo que el inicio no es en realidad más resonante que el N), y si acaso lo que sí que se aprecia es una tendencia, muy clara, a optar por los inicios menos resonantes.

⁶ Recordemos que la Hipótesis TVR la hemos planteado en el capítulo 4 de esta tesis.

⁷ La escala es justamente la inversa de la que planteamos para las propias CS en el capítulo 1.

⁸ En realidad, como $(\text{SON}_{\text{LÍQ}}) > (\text{SON}_{\text{NAS}})$, para tener un $\Delta(\text{SON})$ igual, el inicio de las nasales tendrá que ser menos resonante.

⁹ Los rasgos que definen la sonoridad de un segmento, como es sabido, son los rasgos de voz (*voicing*) y los del modo de articulación. Estos rasgos pueden ir alineados en un plano infrasegmental, y los rasgos (unarios) de LUGAR, en otro plano diferente, dado que el modelo autosegmental tiene la ventaja de disociar los rasgos en distintos planos subsegmentales.

Esto mismo que decimos de los rasgos puede trasladarse a la fonología de elementos; se trata de cuáles son los elementos que tienen en común las CS con las vocales, cuáles se permiten y cuáles no. Elementos que a su vez deben diferenciarse claramente de los propios del inicio.

¹⁰ Por ejemplo, hay que tener en cuenta la influencia de cúmulos consonánticos y no sólo de la consonante inmediatamente precedente, como muy bien advierte Toft (2002), y como analizaremos en el capítulo siguiente.

¹¹ Lo analizaremos con más detalle en el próximo capítulo. Y también veremos otro dato que parece contradecir la norma de que los inicios menos resonantes sean los más favorables: veremos cómo las sibilantes, a pesar de ser fricativas, parecen ser menos vulnerables a la presencia de cúmulos consonánticos que las oclusivas.

¹² Habíamos citado /f/ y /v/, pero también deberían figurar aquí /θ/ y /ð/.

¹³ Nosotros suscribimos la clasificación de Ewen & van der Hulst (2000). Caben otras posibilidades de clasificación, pero la idea de las clases naturales está aceptada de manera general. Y entendemos que resulta útil para tratar la afinidad de los distintos inicios por las CS.

¹⁴ Los elementos fonológicos, en cambio, pueden ser inadecuados para representar el comportamiento de los inicios. No vemos cómo expresar la disociación entre ‘sonoridad’ y ‘lugar de articulación’ que acabamos de reflejar en el diagrama.

Y por otra parte, sólo es fácil expresar un punto de articulación compartido entre inicio y CS (como elemento compartido); pero cuando no es compartido, no vemos cómo reflejar el mayor o menor parecido entre los distintos elementos. ¿Qué hacer entonces? Aunque el modelo de elementos es bastante cualitativo, pensamos que los elementos **A,I,U** le dan también un cierto carácter concéntrico. Nos parece razonable tomar como centro al elemento coronal **T**, y a partir de ahí trasladar los otros elementos de lugar a la escala numérica que hemos propuesto.

¹⁵ El hecho es propio del ‘germánico’: en alemán también podemos encontrar *bleiben* con [m].

¹⁶ La nomenclatura sólo pretende facilitar las cosas, y por eso nuestra propuesta es a título provisional. Si el concepto prospera, podrían plantearse otros términos, tales como ‘**semigeminadas quirales**’, ‘**invertidas**’ o ‘**en espejo**’, por ejemplo.

¹⁷ Recomendamos Roca & Johnson (1999: 543ss.) para una descripción elocuente de lo que son las reglas ordenadas en fonología.

¹⁸ Desde el punto de vista fonético, hay una sola emisión sonora, común para inicio y CS. El aire de la CS nasal sale por la nariz, pero sin perder la articulación de la T del inicio, que mantiene cerrada la salida por la boca.

Nos da la impresión de que hay un solo chorro de aire, que remeda a las **africadas**: primero se articula la oclusiva, y después al relajar la oclusión hay una salida. Y desde el punto de vista fonológico, la nasal y la oclusiva comparten plenamente el rasgo de [lugar] y el rasgo [-cont], en una sola secuencia.

¹⁹ También en la evolución de los seres vivos: para poder adaptarse al medio, hay que tener unos mecanismos suficientes, ya de base. Es la **evolución desde dentro** de Stephen Jay Gould, que explica Sampedro (2002, capítulo 5). Dice Sampedro (2002: 75) que *los graduales oficios del entorno cambiante –la adaptación progresiva y dirigida desde fuera – son menos relevantes que las variaciones, a menudo drásticas, que imponen desde dentro la dinámica intrínseca de los genomas y la lógica genética profunda de la construcción de un ser vivo*. Sampedro vuelve sobre la idea de “evolución impulsada desde dentro” en (2002: 104). Recomendamos con el mayor entusiasmo la lectura de su libro entero, por su rigor y su capacidad explicativa.

²⁰ Aparte de las consonantes que impiden la silabicidad de la /l/ y de las que la favorecen, hay un tercer rasgo que indica que la /l/ tiene cierta sensibilidad al contexto izquierdo y no es del todo inmune, frente a lo que se apuntaba en Toft (2002: 116ss.). Este tercer rasgo es la influencia de cúmulos consonánticos a su izquierda, que Toft sólo había considerado para las nasales, pero que podría afectar también a las líquidas, según veremos en el siguiente capítulo. Si no tratamos aquí este aspecto, es porque nos parece que el factor determinante no es la mera presencia de cúmulos (a un nivel puramente segmental, como sugiere Toft), sino la distribución métrica o silábica que ello acarrea.

²¹ El **Principio del Contorno Obligatorio** (PCO) evita las secuencias de elementos idénticos, tanto si son tonos como si son segmentos. Odden (1986), en Goldsmith (1999) es un referente clásico con respecto a los tonos, y en Harris (1994: 172) se aplica a los segmentos: *At the melodic level, adjacent identical units are disfavoured*. Como dice el propio Harris, el PCO no es exactamente un principio rígido (aunque el nombre se ha quedado así), sino más bien una fuerte tendencia que se observa en el lenguaje humano. (Goldsmith (1995: 21) recoge la opinión de Yip, según la cual el PCO no es un *desencadenante* de reglas, pero sí un *bloqueador* de reglas).

²² Una excepción significativa es *library*, que tiene la variante secundaria [ˈlaɪbrɪɹɪ]. Esta forma de distancia cero es posible gracias a que hay una vocal en el contexto derecho de la CS.

²³ Se puede formalizar la identidad con matrices de rasgos ordenados, que se cotejan de uno en uno. Asignamos valores numéricos a la comparación: si el valor de especificación del rasgo coincide, le asignamos un 1, y si no, un 0. La identidad de dos segmentos es total cuando sólo hay ‘unos’ en la matriz compuesta. Con este mecanismo podemos ‘cuantificar’ el grado de

coincidencia, incluso para rasgos unarios, que dispondremos en el mismo lugar de la matriz, para hacer posible la comparación.

De todas formas, no hay que confiarse por la elegancia de la formalización; es necesario captar la geometría de rasgos, y para ello no basta la simple alineación dentro de una matriz.

²⁴ Se hace necesario entonces postular que existe un nodo σ , algo que la FR siempre ha evitado. Pero la restricción es necesariamente tautosilábica, y por ello no vemos la conveniencia de prescindir de un nodo silábico.

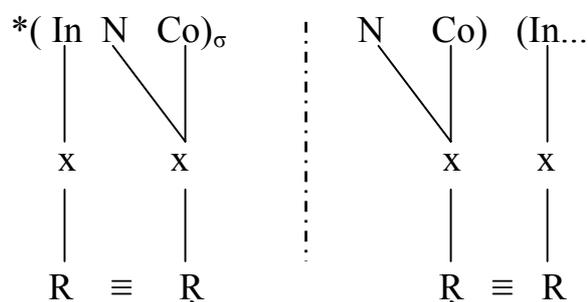
²⁵ Por el mismo patrón, se puede explicar que en palabras del tipo *Trenton* o *wanton* se prohíbe la lenificación total de la /t/ porque se produciría una situación de distancia cero. La hipótesis es que una vez eliminada la /t/, la /n/ de la coda se reasigna al siguiente inicio (esto es, *wanton* → *wanon*), por lo que las dos /n/ pasarían a estar en la misma sílaba.

9 Relación inicio/CS: el plano superior

El capítulo anterior estuvo dedicado a explorar la influencia del contexto izquierdo en la porción más baja de la estructura fonológica, en el plano segmental. Vimos entonces la estrecha relación que puede establecerse entre la CS y su inicio en lo respectivo a los segmentos y a los rasgos que los componen. En este capítulo vamos a “cambiar de escenario”, ascendiendo en la estructura para analizar los detalles del plano superior, tanto en el subplano silábico como en el métrico.

Comenzamos el estudio del subplano silábico recogiendo por un instante un detalle del capítulo anterior. Simplemente recordamos aquí que la ‘distancia cero’ opera en el ámbito de la sílaba:

(1)



En la parte izquierda del diagrama se aprecia cómo la identidad de segmentos está prohibida en un dominio tautosilábico, estrictamente. Por el contrario, no hay ningún impedimento cuando los segmentos

son heterosilábicos: la segunda figura del diagrama representa un par de geminadas ‘verdaderas’, y aún con mayor razón se admitirían unas geminadas ‘falsas’ (con linde de palabra de por medio).

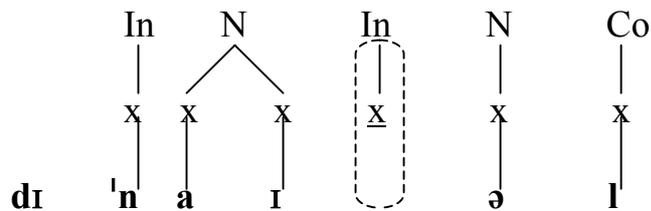
Más sobre el subplano silábico: las CS/ V__

Centramos ahora el estudio del subplano silábico en las CS/ V__, las CS precedidas por una vocal. Como vimos en el capítulo anterior, estas CS están en un contexto izquierdo más resonante que la propia CS, lo cual representa una cierta anomalía. Pero lo cierto es que no se incumple la norma de que el inicio sea menos sonoro que su N, ya que el contexto vocálico no ocupa el inicio de la CS, sino un N previo. Y cuando menos, esa vocal previa proporciona a la CS el respaldo suficiente para cumplir con la condición mínima de tener un contexto izquierdo.

En realidad no se trata de una simple vocal, sino que es todo un diptongo el que precede a la CS. La segunda vocal del diptongo puede ser una /ɪ/ (/aɪ/, /eɪ/, /ɔɪ/, como en *denial*, *trial*, *Lyal*,¹ *betrayal*, *loyal*, o bien una /ʊ/ (la /aʊ/ de *vowel* y hasta de la singular *withdrawal*.² Todas estas CS son débilmente silábicas, ya que tienen una variante abierta, con schwa: (ə)R. En este último caso, tenemos un **triptongo**, (según la nomenclatura tradicional (Cruttenden 2001: 94, 138; *Nueva gramática de la lengua española* (2011: § 8.11h)),³ formado por tres vocales que en realidad son un diptongo seguido de una schwa (2001: 138); por ejemplo, *denial* es /dɪ'naiəl/.

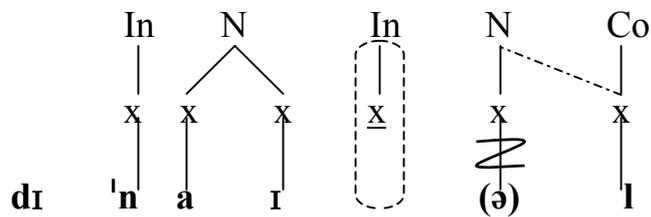
Sin embargo, los análisis más modernos consideran que hay en realidad dos sílabas separadas: en una está el diptongo y en la otra la schwa, de modo que *denial* es /dɪˈnaɪ · əl/:

(2)



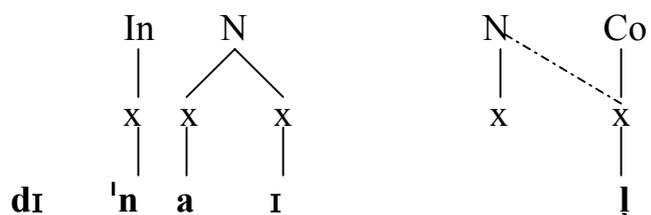
La CS se forma a partir de esta estructura básica, cuando la schwa se reduce hasta cero y la R se liga al N vacío, según el esquema normal de las CS inglesas:

(3)



El diagrama refleja cómo el inicio de la CS está vacío; puede que no haya nodo In siquiera (posibilidad que representamos con la línea discontinua en torno a esa posición). De no haber inicio, tendríamos dos nodos N adyacentes, el del diptongo y el de la propia CS:

(4)



Esta configuración con dos N adyacentes resulta desfavorable, según Harris (1994: 205),⁴ y por ello tiende a evitarse. Se ha formado una CS que tiene contexto izquierdo pero no tiene inicio, en franco contraste con la estrecha relación inicio/CS de otras CS (recordemos las antigeminoides del capítulo anterior, por ejemplo). Ahora hay una distancia estructural entre el diptongo y la CS, que se suma a lo desfavorable de la distancia sonora entre la CS y su contexto izquierdo. Por eso estas CS son **muy débilmente silábicas**, con una alternancia entre schwa y CS en la que deben de predominar claramente las formas con schwa.

Hay ocasiones en que no se forma una CS aun cuando las condiciones parecen ser satisfactorias. Es así en *violet*, *violent* o *violence*, donde la V que sigue a la /l/ debe de ser un factor debilitador de la silabicidad, que ya de por sí es tan débil en este tipo de CS⁵. Tengamos en cuenta que si en estas palabras se formara una CS, el resultado sería una CS/V__V. También pensamos que la falta de silabicidad se debe a un factor cognitivo que interactúa con la fonología: no se forman CS por la dificultad para reconocer ciertas formas. Por ejemplo, *viol* tiene un par mínimo con *vile*. Desarrollaremos este aspecto en el capítulo siguiente, así que no abundamos más en ello por el momento.

Sin salir de la clase de las líquidas, la /r/ forma CS/V__ en *General American* con las mismas pautas que la /l/. Asumimos que se forma una CS tras el diptongo /aɪ/ en *liar*, porque la alternancia entre [ˈlaɪər] y [ˈlaɪr] que da el *Webster* debe de corresponder a /ˈlaɪr/; y lo mismo debe de ser para *fire*, *choir* y *admire*, a pesar de que sólo figuran como

/-aɪr/. Para el diptongo /eɪ/, *mayor* sólo tiene la versión con schwa /^hmeɪ-əɪ/, y si reduce material lo hace suprimiendo la paravocal (esto es, [^hme-əɪ], y también [meəɪ]), en lo que se conoce como **diphthong smoothing** (Cruttenden 2001: 139-40). Con la /ʊ/ como paravocal, *Webster* da como alternante con o sin schwa la pronunciación de *power* y *flower*, pero sólo [^hflaʊr] para *flour*⁶. Otra vez vemos cómo formas casi idénticas tienen comportamientos diferentes, que casi con certeza debemos atribuir a factores externos a la fonología.

Como es de esperar, las nasales /n, m/ lo tienen más difícil que las líquidas para formar CS/ V__ . Son estrictamente no-silábicas *defiant*, *flamboyant*, *lion*, o *ion*. Sorprende la CS en *McMahon* /mæk^hma:(ə)n/, que va precedida por una vocal larga en vez de un diptongo. Y *mayn't* es /meɪnt/, pero tiene una variante secundaria [^hmeɪnt], donde no se ha elidido ninguna /ə/, curiosamente. Con la [-(ə)m], *Webster* da para *diamond* el consabido patrón de alternancia; y *Graham(e)* es no-silábica, según el *Everyman's*.

A tenor de lo visto en todos los ejemplos que hemos citado, las CS/ V__ resultan bastante peculiares. A veces no es fácil percibir la CS, y podemos dudar de hasta qué punto hay tal CS. Pero hay datos que confirman la existencia de esta. Por una parte, al comparar los “triptongos” con sus equivalentes sin schwa. Por ejemplo, la schwa de *denial* puede reducirse y formar una CS; pero nunca llegará a la supresión o ‘cero’ que podemos observar en la forma sin schwa *Nile*, su par mínimo. Y por otra parte, cuando el ‘triptongo’ va seguido de una obstruyente (por ejemplo, en *quiet*, *diaper*, *liable*) no sólo no se

forma una CS, sino que ni siquiera se alterna entre schwa y ‘cero’ (a diferencia de las sonantes). Compárese a tal efecto *diet* con *dial*.

El subplano métrico y la ‘captura de mora’

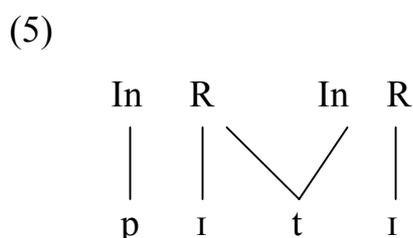
Dejamos atrás el subplano silábico y ascendemos hacia el subplano métrico. Antes de entrar en los aspectos teóricos, queremos introducir la cuestión con un ejemplo ilustrativo. Tomamos la doble acentuación de la palabra *bicentenary*, que el diccionario *Webster* recoge como [bi·cen·'ten·a·ry] y [bi·'cent·n·ar·y] (respetando la notación de este diccionario), y al momento apreciamos cómo al cambiar la acentuación cambia también el silabeo⁷. La diferencia concretamente está en que la /t/ queda **“atrapada”** en la coda previa, mientras que la /n/ se queda sola. Vemos además que esa /n/ da lugar a una CS⁸, y no se nos escapa que la formación de la CS pudiera estar acoplada a la absorción de la /t/.

Cuando estudiamos las CS/V__ en la sección anterior, vimos que se formaban sin inicio. En aquel momento no nos planteamos ninguna generalización, por tratarse de un caso muy concreto y peculiar. Pero ahora con la CS que se forma en *bicentenary* [-¹cent·ŋ-], es ineludible preguntarnos si la ausencia de inicio favorece la formación de CS, en general. Porque entonces tendríamos que conciliar esta ausencia de inicio con la estrecha relación inicio/CS que habíamos apuntado en el capítulo anterior.

Hay otra necesidad de conciliación evidente en el silabeo de *bicentenary*. La absorción de la /t/ no cumple con el requisito de que

las CS estén en una sílaba con inicio, y sin embargo resulta muy natural, o “muy nativo” en inglés. Es porque el silabeo “sin inicio” resulta muy propio de la lengua inglesa, lo cual nos lleva a pensar que se trata de un rasgo general del inglés, más allá del ámbito concreto de las CS. Este hecho hay que referirlo al modelo teórico, que deberá resolver la paradoja de que una lengua conculque el Principio Universal de Maximización de Inicios (*Onset Maximization*) con tanta naturalidad.

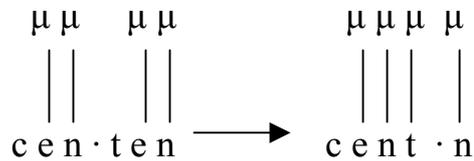
Pues bien, la absorción de la /t/ en *bicentenary* nos recuerda a la **captura de coda** (*coda capture*) que se describe en Harris (1994: 198-202). Dicha captura se enuncia como sigue: *En el ámbito de un pie, conéctese la consonante asociada al inicio de la sílaba átona a la coda de la sílaba tónica.*⁹ Por ejemplo, la captura de coda en *pity* /^hpɪtɪ/ produce esta estructura (1994: 199):



Nuestra sugerencia es reinterpretar esta idea a la luz de la estratificación en planos que venimos trazando en esta tesis. En lugar de una captura de coda, proponemos que se haga una ‘**captura de mora**’, un escalón más arriba. La captura de coda opera con las asociaciones a nodos silábicos, y la de mora opera con asociaciones a nodos moraicos. De este modo, cuando en *bicentenary* pasamos de

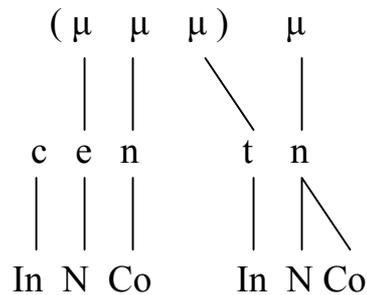
[bi·cen·'ten·a·ry] a [bi·'cent·n·ar·y], la captura de la /t/ se hace estrictamente a nivel moraicó:

(6)



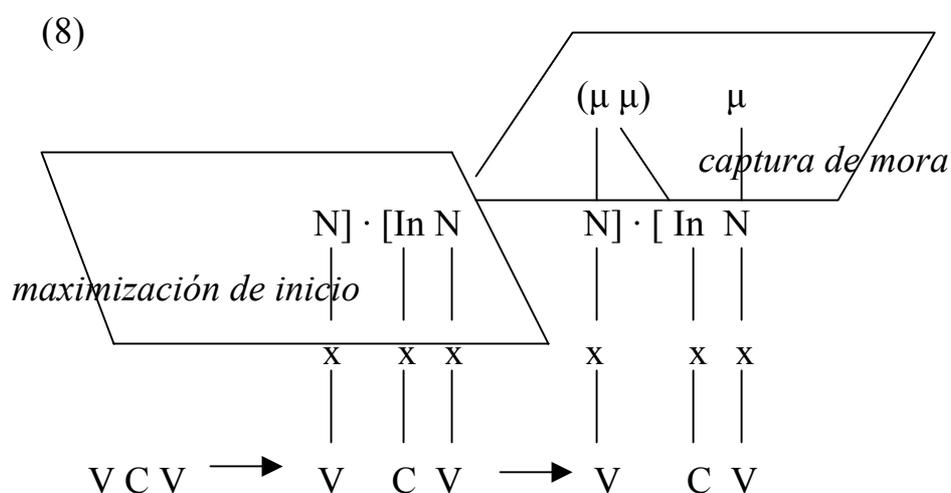
Esto implica que en el nivel silábico las cosas siguen como estaban: la /t/ sigue en el inicio en el que se había silabeado. El panorama completo para [bi·'cent·n·ar·y] es:

(7)



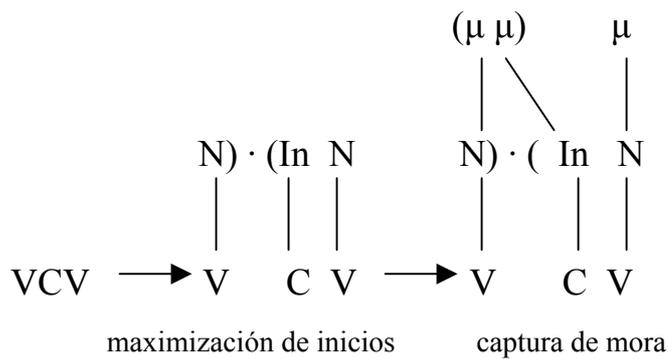
Según este esquema, la /t/ sigue asignada a su sílaba de siempre (σ_2), pero también al pie (P_1) de la sílaba previa. El resultado es que un solo segmento queda a caballo entre dos estructuras, aunque sin llegar a ser ambisilábico. Desde luego que no tiene el tipo de ambisilabicidad que tenía la /t/ de *pity* en el diagrama (5); porque ahora el segmento ya no está asociado a dos nodos de la misma gradilla, y con ello se supera una objeción fundamental a las propuestas de ambisilabicidad que se habían hecho (ver Selkirk 1982: 343, Harris 1994: 198-200, Blevins 1995: 232, 244, y en especial, Harris 2004: §3).

Según la hemos planteado, esta captura de mora nos parece más dinámica que la ambisilabicidad clásica, y quizá pueda ser también más intuitiva. El dinamismo se debe a que está articulada en planos diferentes, de modo que una consonante C puede oscilar entre la asociación a dos vocales distintas, V_1 y V_2 , pero en lugares distintos del almacén estructural:



La captura de mora tiene la ventaja de respetar la Maximización de Inicios. Y a mayor beneficio, la captura de mora también cumple con la condición de **conservar la estructura** (a modo de **Principio de Proyección**) que manda la FR (véase Harris 1994: 189, 193), puesto que no hay resilabeo. Para mayor sencillez gráfica, hacemos ahora un diagrama que refleja paso a paso cómo se cumplen los requisitos escrupulosamente, gracias a que cada asignación ocurre en su lugar correspondiente:

(9)

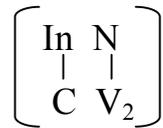


Queda claro entonces que la Maximización de Inicios es una regla silábica, y como tal se aplica en dicho componente de la estructura, mientras que la Captura de Mora se hace en el componente métrico. Damos por hecho que la captura de mora es **opcional**, y en ningún caso anula al silabeo más “canónico” (sin captura), sino que se complementa con él¹⁰. Las dos formas alternan entre sí, de manera que la estructura **resuena** entre la (VC·V) de la captura de mora y la (V·CV) de la maximización de inicios¹¹. Queda así explicada la coexistencia en tensión de dos estructuras que se planteaba al principio de esta sección, o al menos así lo entendemos nosotros.

Obviamente, no todo son ventajas en el esquema que acabamos de proponer. El primer inconveniente es que el inicio pasa a tener peso moraico cuando es capturado. Es un hecho de apariencia extraña, sin duda, y por el momento no podemos hacer nada mejor que asumir que el valor moraico del inicio capturado es equivalente al de una coda. La asimilación entre inicio capturado y coda sugiere que hay algo de resilabeo en la captura, por mucho que se haga en forma moraica. Y damos por hecho que los inicios que no son capturados no tienen peso métrico, como siempre se ha considerado.

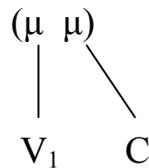
El otro escollo es la falta de alineación entre los dos subplanos que sufre la consonante afectada. Dada una secuencia VCV, la C pertenece a la segunda sílaba (σ_2) por la maximización de inicios, pero en la captura de mora se reasigna al primer pie (P_1), que corresponde al pie de la sílaba previa (σ_1). Por lo tanto, la consonante es **tautosilábica** respecto a V_2 :

(10)



y al mismo tiempo es **'tautopodal'** respecto a V_1 :

(11)



Al estar asociada a dos vocales a la vez (V_1 y V_2), aunque sea en diferentes puntos del armazón estructural, la consonante **oscila** entre las dos asociaciones. De aquí debe de venir la dificultad que tienen los hablantes nativos (e incluso los lingüistas entre ellos) para asignar un silabeo a una secuencia VCV, alternando entre $(V) \cdot (CV)$ y $(VC) \cdot (V)$. Por ejemplo, *pity* alterna entre $(pi \cdot ty)$ y $(pit \cdot y)$. La captura moraica se percibe como 'más nativa' o propia del inglés, al ser el fenómeno específico de la lengua inglesa. En cambio, la maximización del inicio es lo universal, y por eso es el "silabeo" que haría cualquier no-nativo del inglés. Pero no olvidemos que al margen de cuál sea el silabeo que

suene ‘más nativo’, los dos están ahí, y por eso existe la tensión o la oscilación entre uno y otro.

También intuimos que la oscilación por la falta de alineamiento hace que la estructura sea **inestable**. Inestabilidad no implica mala formación, desde luego; es sólo que la estructura es menos estable de lo que sería con una alineación total de sílabas y pies, con el segmento asignado al pie de su sílaba.

Esta propuesta que estamos haciendo sobre los **grados de estabilidad** de las estructuras nos parece novedosa, y también puede resultar útil. Nos parece probable que las estructuras de transición, de baja estabilidad, jueguen un papel importante en los procesos de **cambio lingüístico**. Análogamente a los compuestos químicos, cuanto más inestable sea una estructura fonológica, tanto más **reactiva** será y tanta mayor será su tendencia a **mutar** (hacia una estructura más **estable**).

Las desventajas que acarrea la captura de mora (peso moraico de un inicio más inestabilidad por falta de alineamiento) tienen que estar compensadas entonces por una ventaja mayor. ¿Qué es lo que induce una anomalía de este tipo? Nos inclinamos a pensar que la captura de mora se hace para satisfacer requisitos métricos. Y más concretamente, por la condición de pesadez que se impone sobre las sílabas tónicas en inglés¹². Por eso encontraremos capturas de coda en inglés y posiblemente en alemán, pero no en checo ni en bereber, por ejemplo.

La exigencia de pesadez, por su parte, la atribuimos a un mecanismo cognitivo, destinado a facilitar el reconocimiento o la identificación de palabras. El origen del mecanismo bien pudiera estar en la propensión que tiene el inglés a formar palabras monosilábicas, las cuales son más fáciles de confundir entre sí. A fin de facilitar la identificación de palabras monosilábicas, y de paso ampliar su repertorio, se puede hacer que terminen en consonante¹³: es así como surge la condición de pesadez métrica. No es descabellado concebir que el patrón esté pensado inicialmente para monosílabos pero después se generalice, y se haga entonces extensivo a las demás palabras de la lengua.

Otro mecanismo cognitivo que favorecería el peso moraico en sílabas tónicas es el patrón de luz y sombra (o primer y segundo plano). La parte de sombra es corta y neutralizada (se disminuye), pero la de la luz debe ser pesada (y si no lo es, se aumenta para que lo sea).

En cualquier caso, es coherente con la teoría que venimos defendiendo en esta tesis, donde la cognición induce o motiva la aplicación de reglas fonológicas para obtener un mejor procesamiento de la información.

Captura de mora y CS

Demos por asumido, pues, que en [bi^{cent}·n·ar·y] la /t/ sufre una captura de mora y que la /n/ forma una CS en el núcleo siguiente. El detalle fundamental en este silabeo es que la /t/ se lenifica a [ʔ], y es entonces cuando se forma la [ŋ]. La lenificación de la /t/ a [ʔ]

demuestra inequívocamente que hay una ‘captura’, como explica en gran detalle Harris (1994)¹⁴. En términos más clásicos, la captura lo sería de coda, pero Harris (2004) dice que la lenificación ocurre dentro de un pie; esto implica que la lenificación se hace dentro del plano métrico, y por eso pensamos que nuestra idea sobre la captura de mora es correcta en esencia.

Por cierto que nadie parece haber reparado en una cuestión básica sobre la derivación fonológica. En la fonología de reglas, podía especificarse una lenificación mediante una regla formal, pero con el mapa de la fonología autosegmental es preciso decir *dónde* ocurre. Si la /t/ se lenifica dentro de un pie, como sostiene Harris (2004), entonces el pie tiene que estar ya formado. Por lo cual, no sirven ya los esquemas de lenificación dentro del plano segmental sin más. La lenificación en sí es segmental (o infrasegmental), desde luego, pero están implicados también los planos más altos. En estas circunstancias, hay que dar una explicación más integrada. O bien que el pie actúa sobre el plano segmental, con efecto descendente; o bien que los rasgos suben hasta el plano métrico, y en este caso podría pensarse que la lenificación corresponde a un filtrado de rasgos durante el proceso de ascenso. Una tercera posibilidad sería que subiese todo en forma de estructura escueta (*Bare Phrase Structure*).

Recordamos aquí que esta CS se forma gracias a una articulación combinada entre la oclusiva [ʔ] y la propia nasal, [n], como ya explicamos en el capítulo anterior al hablar de las antigeminoides. La explosión (*release*) de la oclusiva se hace con los labios cerrados, y así es como obtenemos la [ŋ]¹⁵. Recordamos también que esta

articulación no admite la menor realización de schwa en absoluto, de modo que esta [ŋ] es la CS más fuertemente silábica de todas.

Lo que sí que permite es la alternancia de formas, ya que siempre es posible pronunciar la /t/ sin lenificar, y entonces podremos oír la schwa en todo su potencial. Por ejemplo, podemos pronunciar *mountain* como [ˈmaʊn · tʰən] o como [ˈmaʊnʔ · ŋ], a voluntad (nótese el distinto “silabeo” asociado con cada forma, aunque estrictamente sea una captura de mora, y no de coda).

La [ŋ] formada en combinación con [ʔ] es tan fuertemente silábica, de hecho, que no es sensible al resto del contexto izquierdo. En esto se diferencia de otras CS nasales, que dejan de ser silábicas si tienen cúmulos consonánticos a su izquierda, como veremos en la siguiente sección. Pero con la ŋ/ʔ__ no hay sensibilidad al contexto previo.

Es perfectamente posible ver la oclusiva precedida de vocal, como sucede en *written*, *Staten Island*, *Latin*, *satin*, *Britain*, *mutton*, *cotton*, así como *important* en RP ¹⁶. Dicho sea de paso, la lenificación a /ʔ/ en este contexto intervocálico nos parece un buen indicador de que hay captura de mora. Pero aparte del contexto vocálico, la oclusiva también forma cúmulos con las consonantes que la preceden. Hay muchos ejemplos de cúmulos con /n/ en inglés americano, tales como *mountain*, *accountancy*, *sentence*, *Trenton* o *Clinton*, y hasta en *Washington DC*.

También sucede con otros cúmulos. Por ejemplo, con la /r/ en *curtain*, *shorten*, *importance* o *Burton*, y con la /l/ en *golden*, *Dalton*, *Milton*.

Igualmente sentimos que la /p/ en *captain* puede dar /'kæp²ŋ/, por mucho que *Webster* sólo considere /'kæptən/. E incluso en *something* [¹sʌm²m̩] volvemos a ver este mecanismo. Insistimos en la total ausencia de schwa en todas estas formas.

Con todo, CS y lenificación no siempre van de la mano. La [ŋ] sólo se observa cuando la oclusiva (/t/ o /d/) se lenifica a [ʔ], pero no si se lenifica a [r]. Este segmento es conocido como **flap** o **tap**, y es típico verlo entre vocales, como en *better*, *writer*. Este *flap* [r] no parece implicar captura de mora alguna (a diferencia de la [ʔ] de *written* que veíamos hace un par de párrafos), y quizá esta sea la clave de que no dé lugar a una /n/ silábica.

La falta de captura de mora unida al contexto intervocálico no nos dejan más alternativa que considerar al *flap* como un segmento puramente **ambisilábico**. No alcanzamos a decir mucho más al respecto, salvo que las CS líquidas no son vulnerables a su presencia, sino más bien al contrario. Por ejemplo, se forma una CS con el *flap* en palabras tales como *water* o *better* y *little*, *bottle* o *subtle* [¹sʌr̩].

Un caso peculiar de lenificación de la /t/, en grado máximo, es la supresión que se escucha en americano coloquial para *plenty* [¹pleni] o *disappointing* [¹dɪsə'pɔɪnɪŋ]. Pues bien, la /l/ aún tolera la silabicidad en *gentlemen* (especialmente en *ladies and gentlemen*, aunque no es tan claro en *gentle*), y lo hemos oído decir en *mental* [¹men̩]. Sin embargo, no puede formarse en *dental* [*¹den̩], ni tampoco en *rental* o *parental*. Es bastante probable que esta diferencia se deba, como

hemos visto en otros casos, a la posibilidad o no de reconocer la forma original por parte del oyente.

Pero el reconocimiento de la forma original aquí tiene un correlato fonológico, estructural. Tengamos en cuenta que en el paso de *mental* /¹mentəl/ o incluso /¹mentl/ a la forma con supresión /¹menəl/ o /¹menl/ hay también un cambio de silabeo importante. La /n/ que en la forma sin suprimir estaba en la coda de una geminada parcial, ahora pasa a ser el inicio de la sílaba siguiente. Por eso entendemos que la supresión es viable solamente en formas de uso muy frecuente y de cierta extensión, como es el caso de (*ladies and*) *gentlemen*, donde es más fácil reconstruir la forma original a partir del contexto y no se presta a confusión con otras formas análogas.

En definitiva, las distintas lenificaciones de /t/ tienen una distribución específica respecto a las CS. La /r/ y la supresión se asocian a la /l/ exclusivamente, mientras que la /ʔ/ lo hace con la /ŋ/. La estructura fonológica, silábica, probablemente juegue un papel determinante, según venimos defendiendo, y no descartamos que el análisis de estas variantes tenga un alcance más general, para el conjunto de la teoría.

Terminamos esta sección con una variante de la captura de mora. La forma con CS de *family* es /¹fæmli/, pero en inglés coloquial se puede insertar una /b/ para formar *fambly* /¹fæmbli/ (que también se recoge en inglés escrito).¹⁷ La consonante epentética se ha insertado justo antes de la CS, y produce una disimilación acústica, pero además una separación estructural. En efecto, la /b/ insertada se interpone entre la /m/ y la /l/ de *family*, pero además se convierte en el inicio de la sílaba

siguiente, y así deja a la /m/ en la coda de la sílaba previa. Por eso nos preguntamos si es que la /m/ ya estaba sujeta a una captura de mora en /fæmɫɪ/, de manera que ya tenía una distancia estructural con la CS. El argumento no es definitivo, ya que la inserción de la /b/ ha acarreado un cambio de silabeo, pero la sensación de separación métrica entre la CS y la C de su inicio vuelve a estar presente.

Cómo influyen los cúmulos consonánticos

Toft (2002: 136-7) ha observado que la /ŋ/ es vulnerable a la presencia de cúmulos consonánticos a su izquierda. Su estudio se centra en la actividad inhibidora del cúmulo /nt__ , y es cierto que *repentance* y *wanton* son estrictamente no-silábicas¹⁸. Nuestra apreciación es que casi cualquier consonante que forme un cúmulo con la /t/ o la /d/ del inicio bloquea la formación de [ŋ]. La silabicidad es imposible cuando la primera consonante del cúmulo es una oclusiva (*reluctant*, *plankton*). Para las fricativas se mantiene la tendencia inhibitoria, bien produciendo formas no-silábicas (*eastern(most)*, *western*, *capstan*), o si acaso con silabicidad débil (*resistance*, *assistant*, *constant*, *substance*, *sextant*). En cuanto a las líquidas, la /l/ produce silabicidad débil en *molten*, *stilton*, *subaltern*, *embolden*.

Para explicar esta influencia de los cúmulos, nos decantamos por ir más allá de las meras secuencias lineales de segmentos. Ya en Hammond (1999: 202) se habla de un cambio de silabeo, pero nosotros lo situamos un poco más arriba, en el subplano métrico. Por ejemplo, en un cúmulo con oclusiva como el de *plankton*, el cúmulo produce una separación muy nítida de las dos sílabas, y tanto es así

que cada una se aloja en un pie independiente. De modo que no se forma un troqueo, y sin troqueo no se puede formar una CS, al menos en inglés:

(12)

$$\begin{array}{cc} \left[\begin{array}{c} M \\ \mu \mu \end{array} \right] & \left[\begin{array}{c} M \\ \mu\mu \end{array} \right] \\ (plæŋk) & (tən) \end{array}$$

El mérito de la captura de mora está en lograr lo contrario de lo que hace el cúmulo. Si el cúmulo separa, la captura une. La consonante capturada está asociada a los dos núcleos silábicos de un troqueo (aunque en subplanos diferentes) y así establece una conexión más fuerte entre ellas. La consonante **resuena** entre las dos sílabas del troqueo, y así lo articula mejor: entonces el N₁ queda más cerca del N₂ y tiene así más fuerza para licenciarlo¹⁹. Además, al ser capturada la consonante, se refuerza la estrategia de dar más peso a la parte fuerte del troqueo en detrimento de la parte débil: así el contraste entre fuerte y débil es mayor, y el procesamiento métrico es más eficiente.

A esto hay que añadir que el grado de captura no es el mismo para todas las consonantes afectadas, sino que hay todo un espectro, con toda probabilidad. Así que no es un fenómeno **todo o nada**, sino que se puede cuantificar mediante porcentajes. El porcentaje mide el grado de captura que sufre una consonante, y por tanto su grado de resonancia entre sílabas²⁰. No es necesario aclarar que cuanto más alto sea el porcentaje de captura de mora, tanto mayor será también el grado de silabicidad de la CS que se forme. Incluso en presencia de cúmulos, podemos comparar, y así vemos cómo en *mention* el grado

de captura es alto, y por ello la /n/ es bastante silábica. Para *resistance* hay algo menos de captura, en cambio, y produce una silabicidad más débil. Finalmente, *wanton* tiene un grado muy bajo de captura y por eso la forma es no-silábica.

Los distintos grados de captura se aprecian más aún con un par ultramínimo: traemos a colación de nuevo las dos maneras de pronunciar una misma palabra, *mountain*, y recordamos cómo se producían “silabeos” diferentes, con captura de mora y sin ella. Cuando se pronuncia [ˈmaʊn · tʰən], la aspiración de la /t/ indica que es cabeza de su pie, por lo cual se han formado pies distintos y no hay troqueo ni CS. Por el contrario, en [ˈmaʊnʔ · ɳ] la lenificación de la /t/ es síntoma de que ha sido capturada, y de ahí que se forme una CS en toda regla.

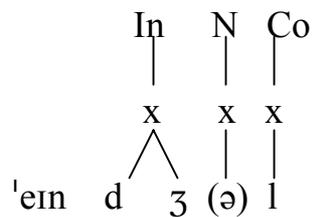
La presencia de cúmulos consonánticos en el contexto izquierdo produce un efecto curioso. En ausencia de cúmulos, habíamos visto en el capítulo anterior que las oclusivas eran más favorables que las fricativas (por Δ SON). Sin embargo, cuando se forman cúmulos son más sensibles las oclusivas /t/ y /d/ que las fricativas sibilantes. El hecho es que los cúmulos no parecen afectar al contexto /{s, z}__ y quizá tampoco al contexto /{ʃ, ʒ}__.²¹

Nuestra hipótesis es que las sibilantes son más proclives a la captura de mora, aunque también es posible que influya un cierto grado de ‘transparencia’ de estas consonantes, que facilitaría la formación de cúmulos sin que se afecte demasiado la estructura general. *Everyman’s* da como fuertemente silábicas las monomorfémicas

mansion, sentient, crimson, o vixen, y asimismo las palabras sufijadas con #-sion# y #-tion# (tales como *mention, declension, function, complexion, option, convulsion*). Son débilmente silábicas *absent* y *ancient*. Parece claro que *absent* es [ab·sent], y nunca jamás [abs·ent], y también *ancient* debe ser [an·ʃent] pero no [anʃ·ent]. Por su parte, *mansion* ya no es tan evidente, y deja margen para pensar.

Un detalle peculiar es que las africadas siempre forman CS débilmente silábicas, tanto con -(ə)n (*question, dungeon, luncheon, oxygen, detergent, agent, legend*) como con -(ə)l (*angel, Michelmores, Rachel y Vachel(l)*). Una vez más, la mera adyacencia lineal de segmentos no explica nada, y el factor determinante vuelve a ser la estructura. La africadas son cúmulos de (oclusiva + sibilante), sí, pero tienen la peculiaridad de que los dos segmentos están en la misma unidad de tiempo (*timing unit*) (Roca & Johnson 1999: 205-9 explica muy bien este aspecto de la Fonología Autosegmental). Por ejemplo, *angel* es /'eɪndʒ(ə)l/:

(13)



Veamos un par totalmente mínimo. En *Hodgson*, la pronunciación estándar es ['hɒdʒsɒn], con una CS fuertemente silábica, ya que la sibilante previa es un segmento simple, en una sola unidad de tiempo. Sin embargo, en el norte de Inglaterra es ['hɒdʒ(ə)n]), débilmente

silábica por la africada que precede a la schwa. Concluimos así que una misma secuencia produce CS fuertemente silábicas cuando la estructura previa no es ramificada, pero sólo las hace débilmente silábicas cuando están en una estructura ramificada, con una sola unidad de tiempo.

La [ɪ] parece ser menos sensible a la influencia de cúmulos previos, pero no del todo insensible. Por ejemplo, el *Everyman's* da *scalpel* como débilmente silábica, a pesar del inicio previo favorable. Y en *sulcal* la variante principal es no-silábica, quizá por tratarse de una palabra infrecuente, aunque también tiene una variante secundaria con CS, seguramente por tener un contexto previo favorable. También la escasa frecuencia de *cobalt* (al tratarse de una palabra técnica, de poco uso) la lleva a ser no-silábica aun sin tener cúmulos:

(14)

$$\begin{array}{cc} \left[\begin{array}{c} M \\ \mu \mu \end{array} \right] & \left[\begin{array}{c} M \\ \mu \mu \end{array} \right] \\ 'kou & bɔlt \end{array}$$

N.B.: La vocal de la segunda sílaba no está reducida a schwa, sino que es una vocal plena. Forma su propio pie, independiente del anterior.

El factor cognitivo condiciona la estructura métrica: distribuye las dos sílabas en pies distintos, independientes, y es por lo que no puede formarse una CS²². Así que esta sección reafirma la importancia de la estructura (fundamentalmente la estructura métrica, de pies) frente a las meras secuencias lineales de segmentos.

Conclusión

Lo importante y fundamental en todo este modelo es que todo está articulado por ámbitos. Cuando vemos que suceden cosas distintas, y a menudo contradictorias, es porque ocurren en ámbitos distintos. Sólo hace falta saber en cuál de ellos está ocurriendo cada cosa, y a ser posible, también por qué. El porqué suele estar fuera del componente fonológico (lo hemos visto en la parte cognitiva, que induce procesos fonológicos para facilitar el procesamiento). Pero si el componente fonológico es capaz de responder a los requerimientos del componente cognitivo, es precisamente por lo bien articulado que está, con tantos ámbitos, subplanos, niveles y demás. Y además podemos tener varias cosas a la vez, porque no hay problemas de “coexistencia pacífica”. Al revés, todo resuena y es polifónico.

¹ Pero *I'll* es sólo [aɪl], según el *Everyman's*.

² En *withdrawal* se puede postular un diptongo /aʊ/ subyacente, que luego da /ɔ:/ en la forma derivada. También puede revelarse útil el análisis con elementos, porque /ɔ:/ está formada por A y U, igual que el diptongo /aʊ/.

³ En teorías menos tradicionales no se habla de triptongos: o bien hay un diptongo seguido de schwa (p.ej., *fire*), como decimos, o bien una de las (para)vocales está en la rama derecha del inicio, y no en el N (p.ej., *music*) (Harris 1994: 60-62).

⁴ *The avoidance of consecutive nuclei is strongly reminiscent of the avoidance of consecutive identical melodic expressions that is embodied in the OCP (36)*. N.B.: OCP es el *Obligatory Contour Principle*, el Principio del Contorno Obligatorio.

⁵ Como veremos en el capítulo siguiente, dedicado al contexto derecho, las vocales /__V son un factor que debilita la silabicidad de las CS, de manera general. Con mayor motivo en una CS que ya es débil de por sí, como son las CS/ V__.

⁶ La diferencia entre *flower* y *flour* representa otro caso de par absolutamente mínimo.

⁷ Queremos resaltar esta relación entre patrones acentuales y esquemas de silabeo en inglés, que nos parece tan interesante.

⁸ *Webster* lo transcribe como pérdida de la /e/ que en el otro silabeo era la vocal tónica.

⁹ “**Coda Capture:** *Within a foot, attach the consonant associated with the onset of the unstressed syllable to the coda of the stressed syllable.*” (1994: 199)

¹⁰ No es de extrañar que la captura de mora sea opcional. Es posterior o subsiguiente al silabeo, no sólo en el espacio, sino en el orden de reglas. También requiere de la estructura de pies para formarse.

¹¹ Así es como *master* puede ser [mas·ter] o [ma·ster], *revels* es [re·vels] o [rev·els], y *pastry* admite [pas·try], [pa·stry], o incluso [past·ry]. (Datos tomados de Ewen & van der Hulst 2000: 120). *It is certainly not the case that native speakers agree on the exact location of syllable boundaries in all the words of a language.* (2000: 120). Conviene aclarar que en algunos casos el silabeo sí que es nítido: *Most speakers of English, for example, will syllabify the words in (10) as shown: [al·ba·tross], [A·me·ri·ca], [slen·der].* (2000: 120).

¹² Este requisito es absoluto en las sílabas tónicas (véase Roca & Johnson 1999: 352-364), pero no es exclusivo de ellas, a juzgar por los silabeos que dan el *Oxford* o el *Webster*.

¹³ Otra manera de conseguir este resultado consiste en asignar tonos, como se hace en mandarín o en vietnamita.

¹⁴ Harris (1994: 194-205) considera la /ʔ/ como un tipo de /t/ lenificada (que puede representarse como /t/), y en general las /t/ lenificadas están asociadas al final de sílaba. Puesto que la /ʔ/ de *mountain* no parece muy diferente de la /t/ implosiva que hay en *mount*, podemos concluir sin mayor riesgo que la /ʔ/ de [ˈmaʊnʔ] está en una coda.

¹⁵ Nuestra impresión es que lo que estamos representando como /ʔ/ consta en realidad de dos componentes diferenciados. En la coda de la sílaba previa queda la /t/, como hemos visto, mientras que el componente glótico /ʔ/ es el **ataque** (o *ictus*, que no es lo mismo que el Inicio) de la propia consonante silábica, análogo al que tienen las vocales en *be²arbeiten* o *ver²irren* en alemán. Dicho ataque glótico está en la misma unidad de tiempo que la vocal o la CS, sin llegar a cubrir una unidad de tiempo (*timing unit*) o un segmento entero. Sea como sea, la /ʔ/ es diferente de la /t/, y una y otra están en sílabas distintas.

¹⁶ Es plausible que haya también un ataque glótico /ʔ/, aunque de menor intensidad, en /d___, (*Ferdinand, Sheridan*) y en /rd___ (*burden, pardon, garden, Gordon, Jordan, Arden*), si bien *Webster* lo descarta para /nd___ (*tendon, abundance, dependency* y *ascendant*). En contexto intervocálico, la silábica de *trident* puede achacarse a la /ʔ/.

¹⁷ La disimilación de *fambly* por *family* aparece citada en Aitchison (2001:160). Y el buscador de internet Google da 18400 referencias con *fambly*, que se interpreta como una forma coloquial de decir *family*; por ejemplo, *fambly album*.

La inserción puede lexicalizarse, como se hizo en *bramble* < *brēmel* < *braemel*. Si es cierto que la inserción de la /b/ se hizo para evitar la contigüidad de nasal+líquida, como sostiene Aitchison (2001: 160) (a partir de Ohala (1974a), que da también *ambrosia* < *amrotia* en griego), entonces hay que inferir que *brēmel* era /^hbre:m|/, ya silábica. Y parece claro que *humble* < *humilis*, igual que *assemble* < *assembler* < *ad* + *simul* tuvo que seguir el mismo proceso.

¹⁸ Preferimos estos nombres comunes en lugar del nombre propio *Kenton*, el único ejemplo que Toft (2002: 136) menciona en su análisis.

¹⁹ Estamos asumiendo que el núcleo tónico es quien licencia el paso de V a schwa en la parte débil del troqueo. La vocal reducida depende de la vocal plena y tónica, queda en relación con ella.

²⁰ Recordemos que ya habíamos hablado de resonancia y porcentajes al hablar de los grados de silabicidad (fuerte o débil) de las CS en el capítulo 2.

²¹ Los datos favorables provienen del *Everyman's*. Sin embargo, en el *Webster* las sibilantes palatales, /ʃ/ y /ʒ/, no son contexto para /l/ ni para /ŋ/. Este diccionario ni siquiera acepta la silabicidad en sibilantes alveolares, porque da como no-silábicas *Nelson, Wilson, Vincent, Jackson, Hudson, Judson*.

²² El mecanismo métrico está en el polo cognitivo de la derivación. En las palabras de uso frecuente es característico que haya agrupación de las sílabas en un solo pie trocaico, reducción de vocal y captura de mora con formación de CS. Todo ello va destinado a facilitar el procesamiento, según venimos diciendo. Cuanto más uso tenga una palabra, más se implementará la reducción vocálica y la captura de mora.

Por eso mismo vemos dos pies independientes en una palabra técnica como *cobalt*, sin posibilidad de captura de mora ni de reducción a schwa de la vocal átona.



10 Contexto derecho

Después de los dos capítulos dedicados al contexto izquierdo, llega el turno del contexto derecho. En ese lado de la CS hay varias opciones a elegir; ya conocemos el contexto absoluto [CS / __#] en las CS a final de palabra, que es el más típico (*cotton, little*). Pero también puede haber una consonante [CS / __C], bien en coda compleja (*thousand, present*) o en interior de palabra (*mistletoe, messenger*). Y también es posible una vocal [CS / __V], que puede verse tanto en interior de palabra (*platinum, Italy*) como tras la adición de sufijos (*humbly, softening*).

Vamos a agrupar las dos primeras opciones en una clase natural, {C, #}, habida cuenta de que el grado de silabicidad es básicamente el mismo en los contextos consonántico y absoluto, CS/ __{#, C}. En cambio, el contexto vocálico CS/ __V produce un claro debilitamiento de la silabicidad, y por eso lo veremos en otro apartado. En definitiva, vamos a estudiar del contexto derecho ateniéndonos a la oposición clásica entre {C, #} y V.¹

Los contextos consonántico y absoluto

El contexto absoluto, CS/___#, implica que la CS está al final de la palabra. Pero entre las CS que van seguidas por una consonante, CS/___C, caben tanto las de mitad de palabra como las de final. La diferencia no es trivial, en tanto que el grado de silabicidad no debe de ser exactamente igual en cada una de las CS, como pasamos a comprobar en el siguiente análisis.

Es obvio que las CS de mitad de palabra favorecen la estrategia métrica. Dada una palabra trisilábica, con tres núcleos (N_1 , N_2 , N_3), la CS coincide con el segundo; como los acentos recaen sobre N_1 y N_3 , el N_2 resulta aún más átono, en proporción. Por ejemplo, en *mistletoe*:

$$(1) \quad \begin{array}{c} \left[\begin{array}{ccc} M & & M \\ \mu & \mu & \mu \end{array} \right] \\ \text{mis tle toe} \end{array}$$

Este aligeramiento que sufre el N_2 (en comparación con N_1 y N_3) favorece la estrategia cognitiva de poner a la sílaba de la CS en un segundo plano, para facilitar el procesamiento fonológico y cognitivo. Recordemos que esta tesis empezó por asumir que las CS responden a una estrategia cognitiva que se vale de un mecanismo métrico. Pero hay más; al quedar el N_2 algo en penumbra, la situación propicia una **conflación de línea**, a lo Hayes², y esto favorece aún más la computación. La conflación de línea supone que el mecanismo de computación tiende a ignorar la línea basal (o **línea 0**), que es la de las moras simples (μ), y opera más directamente con la línea que está

encima, la de las supermoras (M). En esas condiciones, sólo hay que atender a dos elementos, y con ellos se forma un ‘**supertroqueo**’, sin más complicación:

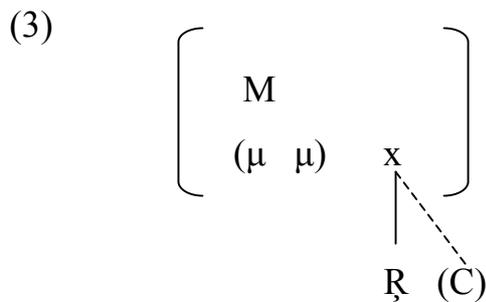
$$(2) \quad \begin{array}{c} \left[\begin{array}{cc} \mathbf{M} & \\ \mathbf{M} & \mathbf{M} \end{array} \right] \\ \mu \quad \mu \quad \mu \\ \text{mis tle toe} \end{array}$$

N.B.: más que de confluencia de línea, nosotros preferimos hablar de un ‘**difuminado de línea**’, según se plasma en las moras de nuestro diagrama, que sólo están menos marcadas pero no borradas. El difuminado implica que el mecanismo de computación no presta tanta atención a ese nivel (difuminado), y opera más con el nivel más señalado; pero no se llega a borrar nada, porque lo difuminado sigue estando ahí. Es sólo una cuestión de grados de atención, donde unos elementos quedan resaltados en primer plano y los otros, más difusos en segundo plano.

Recapitulamos diciendo que cuando la CS queda “emparedada” entre dos acentos, su silabicidad se favorece: al disminuir su peso relativo, es más favorable desde el punto de vista métrico. Traducido a ejemplos concretos, el *Everyman’s* da *falcon* como no-silábica, pero *falconry* y *Falconbridge* como débilmente silábicas³. El proceso es igual entre *open* (que no es silábica) y *openness* (que sí tiene CS, según *The Advanced Learner’s Dictionary of Current English*). De modo que el *sándwich* de acentos, lejos de inhibir a la sílaba de la CS,

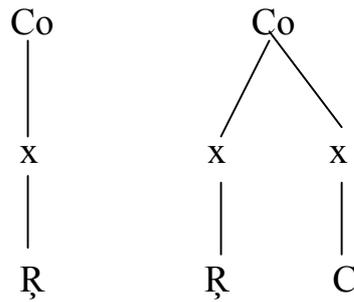
es un contexto prácticamente ideal para formar CS, quizá el más favorable de todos⁴.

Las CS de final de palabra, tanto CS/ __# como CS/ __C# , se limitan a ocupar la parte átona del troqueo, sin el aligeramiento extra del *sándwich* de acentos. La cuestión para las CS/ __C# es si la presencia de una segunda consonante las hace un poco más pesadas, dejándolas menos en ‘sombra’ o ‘segundo plano’. No debería ser así, porque la proyección métrica no debe distinguir entre una consonante o dos. En nuestro modelo decimos que la coda compleja se filtra en su ascenso al nivel moraicó; que la derivación es semitransparente, buscando sólo aportar ‘cantidad suficiente’ al troqueo que se forma. Es lo que ya hemos sostenido al hablar de la derivación métrica, y por eso deducimos que la silabicidad de la CS no debería verse afectada en este aspecto:



Sin embargo, nuestro modelo tiene en consideración el conjunto de la estructura. A pesar de la igualdad en el tratamiento métrico, el plano silábico sigue estando presente, y allí sí que hay una leve diferencia entre coda simple y coda compleja:

(4)

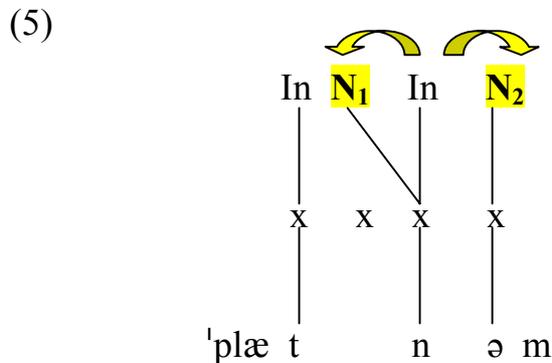


Pero aquí la “pesadez” de la coda no juega en contra, sino a favor de la silabicidad de la CS. Nos da la sensación de que la segunda consonante puede “empujar” a la CS aún más hacia el N. Pensemos por ejemplo en *open* y *opened*, donde el grado de silabicidad es equivalente o aun mayor para esta segunda. Por cierto que los pares mínimos como este indican que las lindes de sufijo no afectan a la estructura silábica, al menos en lo relativo a las CS. Por ejemplo, al añadir un sufijo flexivo de tercera persona {-s}, o el de pasado {-ed} (*tackle(d)*, *whistle(d)*, o *button(ed)*).

El contexto vocálico

Pasamos a la otra clase natural. La diferencia entre los dos tipos de contexto (esto es, {C, #} versus V), es clara. El contexto vocálico tiende a debilitar la silabicidad de la CS, incluso hasta el punto de anularla. Es fácil ver la menor silabicidad de las CS/ __V en *opener* respecto a *open*. También se siente la influencia de la vocal en las CS de interior de palabra, en las monomorfémicas *platinum*, *retina* igual que en la derivada *vaccination* ⁵.

En nuestro criterio, la menor silabicidad (debilitamiento o anulación) por contexto vocálico se debe a la estructura. Para explicar mejor este punto, dibujamos la CS/ __V en un diagrama:

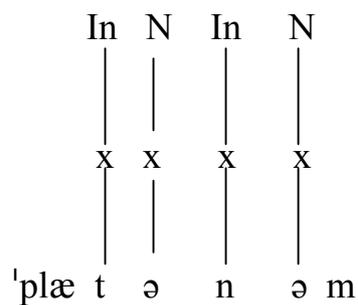


Ahora vemos cómo la CS/ __V se debate entre dos núcleos silábicos, que diremos que funcionan como ‘**atractores**’. El N₂ es el que absorbe a la CS en su esfera de influencia, por el Principio de Maximización de Inicios. Por lo cual la CS resulta ser **heterosilábica** respecto al N₁, que es el que la hace ‘silábica’.

En esto vemos por qué se atenúa la silabicidad. En el tipo habitual de CS (/ __{C, #}), la R está en una coda, dentro de la misma sílaba del N hacia el que ramifica. Sin embargo, en el contexto / __V la CS forma el inicio de esa vocal. Pero como además sigue ramificando hacia el N previo (que está en otra sílaba), queda “**sobrecargada**”. Podemos decir en toda regla que la CS/ __V es un segmento **ambisilábico**. Es una configuración peculiar. Al vivir a caballo entre dos sílabas, y no estar alineada con una única sílaba, la CS pierde algo de su silabicidad, que puede llegar al extremo de anularse. Y a buen seguro que una estructura así es **inestable**.

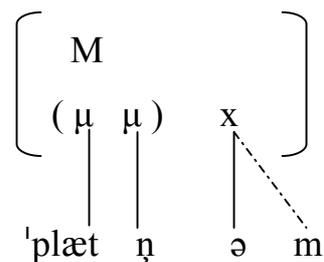
Para ser más precisos, y de paso no vulnerar las reglas, volvemos a plantear la ambisilabicidad en términos de **captura de mora**, el concepto que ya propusimos para el contexto izquierdo. La idea es que en el plano silábico la ‘R’ sólo está asociada a un inicio, sin mayor problema porque ahí la schwa se expresa plenamente:

(6)



Es en el ascenso al plano métrico cuando se reduce; la schwa se filtra hasta el punto de que se forma una CS, como de costumbre, por captura de mora:

(7)

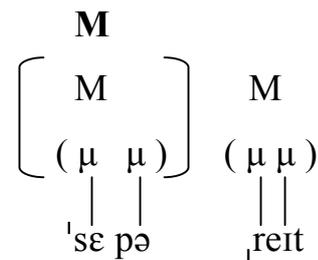


Una vez más nos resulta útil tener en cuenta la proyección métrica, y postular que las CS inglesas se forman en ese nivel. También así podremos entender cómo una misma palabra (*separate*) tiene dos realizaciones distintas, como destaca Harris (1994: 183)⁶. Para nosotros, se trata de un caso más de derivación gramatical (verbo

oxítono; nombre y adjetivo paroxítono) tan propio del inglés. Así se hace en *torMENT* (verbo) / *TORment* (nombre) y en *diGEST* (verbo) / *DIgest* (nombre) [Roca & Johnson 1999: 295-6 da una lista bastante exhaustiva de ejemplos]. En esta misma línea, el verbo *séparàte* /'sepə,reit/ es estrictamente no-silábico, mientras que el adjetivo, *séparate* /'sep(ə)rət/ muestra una alternancia entre schwa y CS. Partimos de este punto para dar lugar a estructuras métricas diferenciadas.

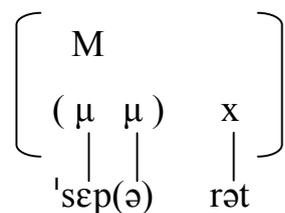
En el verbo se forman dos pies independientes. La palabra consta de dos troqueos, de los cuales el primero es cabeza de la palabra, y en consecuencia recibirá también el acento principal (que representamos en el diagrama con una **M**). Pero sobre todo, la sílaba final tiene acento secundario y la vocal es plena (un diptongo). La proyección en su conjunto es:

(8)



En cambio, el adjetivo sólo tiene un troqueo⁷, y por eso vemos una vocal neutra en /'sep(ə)rət/, que además puede reducirse a CS (Harris lo llama *síncope*). Al no haber más que un pie, no puede haber acento secundario; por eso la sílaba final lleva la vocal reducida, y en vez de una mora le asignamos un aspa, como ‘resto’ del troqueo:

(9)



Como indica el diagrama, la schwa puede perderse y dar lugar a una CS, que formaría la segunda mora de la supermora.

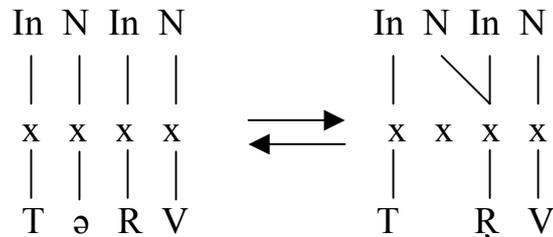
Opciones del contexto vocálico

El mecanismo que hemos postulado nos permitirá explicar las opciones del contexto vocálico, que no es tan sencillo como el consonántico por las variantes que tiene. Pasemos al análisis. Cuando abordamos el estudio de las CS, planteamos una gradación de su silabicidad. El espectro se correspondía con dos variantes del proceso de reducción, la forma con schwa (o abierta) y la forma con CS. Si sólo existe forma con schwa, decíamos que es no-silábica, y cuando no hay nada de schwa hablábamos de CS fuertemente silábicas. Entre medias están las CS **débilmente silábicas**, en las que se observa una alternancia entre schwa y CS. Las CS/___{#,C} se amoldan bien al concepto de silabicidad débil, pero las CS/___V tienen un espectro de realización que abarca hasta cuatro opciones. Por eso vamos a dedicar esta sección a verlas en detalle.

PRIMERA OPCIÓN: Es el primer paso de la reducción: la forma con **schwa** alterna con una CS. Son las CS débilmente silábicas de

siempre, y así es como oscilan *questioning* [-tʃənɪŋ/-tʃɪŋ] y *receptionist* [-ʃənɪst/-ʃɪst].

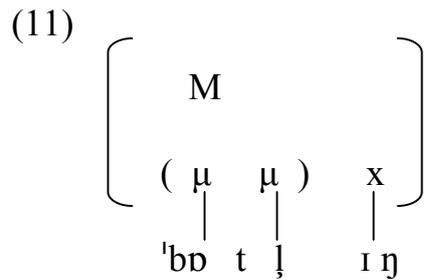
(10)



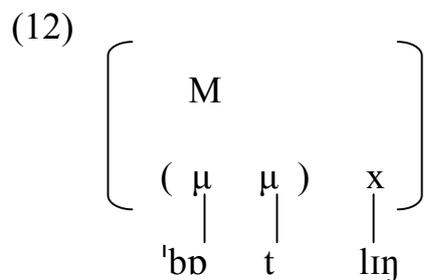
Eso sí, ya apuntamos que el contexto vocálico añade a las CS la complicación de un posible carácter “ambisilábico”. Y este es sólo el principio de las dificultades.

SEGUNDA OPCIÓN: Se avanza otro paso más en el espectro de reducción: la CS alterna con la ‘supresión’, que supone el más alto grado de reducción. Por ejemplo, *bottling* puede ser tanto [ˈbɒt.lɪŋ] como [ˈbɒt.lɪŋ]. Hemos marcado la estructura de sílabas para resaltar cómo en la variante suprimida se pierde una sílaba respecto a la de la CS.

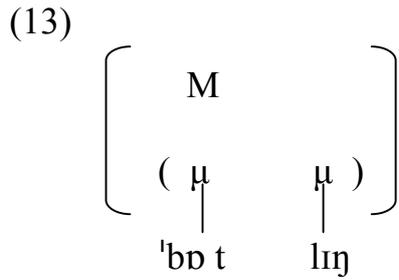
La forma con CS, [ˈbɒt.lɪŋ], tiene la misma estructura que propusimos para *separate* en su forma “sincopada”, cuando es un adjetivo. Hay un troqueo, con dos moras agrupadas para formar la supermora (M), más el resto (x) que la complementa:



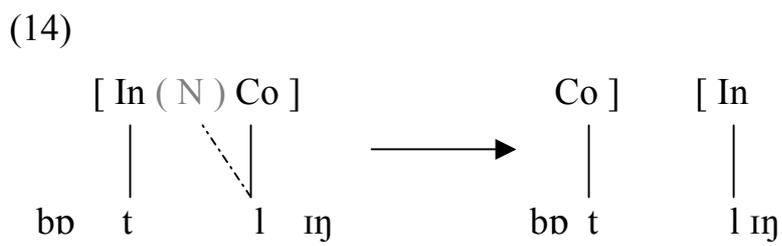
Así tenemos distribuidas tres sílabas en una estructura binaria. Si a ello le sumamos el carácter ambisilábico de la CS, poco puede extrañarnos que fuera inestable. Esta estructura tiende a **mutar**. De modo que la reducción viene a dar un paso más, para mejorar la estructura métrica en *bottling* como lo hacía en *separate*. Porque si la CS dejara de ser silábica, dejaría de estar a caballo entre dos sílabas, con lo que mejoraría su alineamiento. Por eso se acomete la supresión, desde el mismo plano métrico, y se forma un troqueo más canónico. Una opción es aprovechar la ‘pesadez’ que aporta la otra consonante:



La /l/ ha salido de la órbita de la captura de mora, que ahora se le asigna a la /t/. Como ya tenemos una coda para capturar (que es la /t/), la /l/ puede quedarse en su lugar de origen, funcionando como inicio de la sílaba siguiente⁸, y así no queda a caballo entre dos sílabas. Ya por sí sola, esta nueva estructura debería ser más **estable**, y aún puede avanzar en la reducción con este fin:



Es lo mismo que se observaría en *separate* como /¹seprɪt/. La estabilidad que aporta la supresión es impecable desde el punto de vista métrico. Pero a cambio hay que contar con alguna dificultad en el plano silábico, porque nos hemos dejado toda una sílaba por el camino. ¿Cómo se hace esto? Pensamos que la CS se desvincula de su N, y entonces el nodo mismo se pierde:



Hemos marcado el lugar de la CS como “coda” para que se entienda su asociación con el N previo. Pero una vez liberada de ello, queda como inicio puro de la sílaba siguiente, según vemos en la segunda parte del diagrama. Y la /t/ deja de ser inicio para reasignarse como coda de la sílaba previa. La /t/ y la /l/ ahora carecen de relación fonotáctica entre ellas, por lo que se ha formado un **pseudocúmulo** (*bogus consonant cluster*) como los de *atlas*, *athlete*, *atmosphere*, *kidney* (véase Harris 1994: 67, 222-3)⁹.

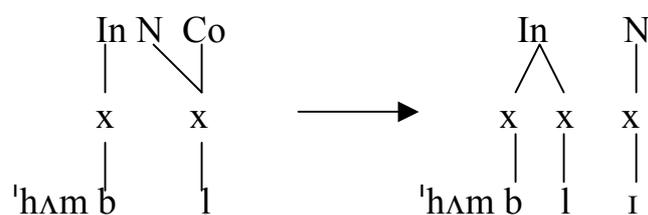
A tenor de lo que estamos viendo, parece que el **resilabeo** es parte corriente de la manera de proceder del sistema; bien al contrario de un

postulado básico de la FR. Pero la mayor dificultad teórica está en que el procesamiento tenga algo de “**descendente**”. Por ejemplo, cuando la CS se desliga de su N en la supresión, tendrá que hacerse desde el plano métrico hacia el plano silábico. El N tiene que eliminarse con la estructura métrica ya formada, porque si no jamás habría podido formarse una CS. Este proceder no es de por sí incorrecto, pero plantea interrogantes sobre el sentido de las operaciones fonológicas (en general).

Otra complicación en los procesos que llevamos vistos es que hay algunas palabras que tienen hasta **triple alternancia**: una forma con CS, otra con schwa, y una tercera con supresión¹⁰. Por ejemplo, *mention* es sólo /^hmenʃn̩/, pero *mentioning* es /^hmenʃn̩ŋ/ [-ʃn̩ŋ], [-ʃən̩ŋ]; y asimismo, *chancellor* es /^htʃɑ:nsələ/ [-s̩lə], [-slə].¹¹ Como veremos en el siguiente apartado, hay preferencias de unas formas sobre otras, y además los motivos no son exclusivamente formales o estructurales, sino también cognitivos.

TERCERA OPCIÓN: La consonante líquida de la CS se incorpora a un inicio ramificado, que tiene como núcleo a la V del contexto derecho. Por ejemplo, cuando *humble* da *humbly* (o también *humbler* y *humblest*) la /l/ del **sufijo** {-ble}, que estaba asociada a un núcleo¹², se reasigna a la rama derecha de un inicio ramificado:

(15)



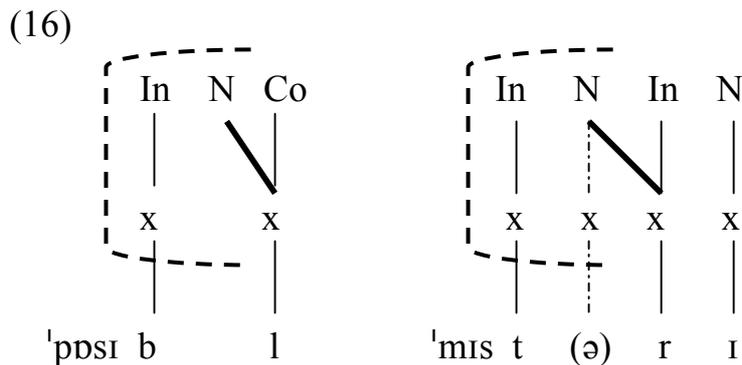
Es interesante ver cómo el sufijo por sí solo no hace que se forme el inicio ramificado, puesto que no lo vemos en *sharp·ly* o *stark·ly*. Pero es un hecho que sí que está en *possibly*. Por mucho que sea difícil obtener pruebas espectrográficas objetivas¹³, hay indicios a favor del inicio ramificado. Uno es que *possibly* tiene que ser siempre [-blɪ], en estricta adyacencia de la /b/ y la /l/; la imposibilidad de la CS (*[-blɪ]) y de la schwa (*[-bəlɪ]) nos demuestra que no hay un núcleo interpuesto (pese a que sí que lo había en origen, porque en {-ble} está la CS, [-blɪ]).

El otro indicio lo encontramos en palabras monomorfémicas como *interesting*, *mystery* o *factory*, que tienen las tres variantes: con schwa plena, con CS, y finalmente, con supresión. En este último caso, la /t/ se articula **distendida** (*released*) y la /r/, ‘aspirada’, en una influencia mutua que Harris (1994: 23) considera privativa de los cúmulos homosilábicos (y que se aprecia al comparar los cúmulos de *petrol* o *nitrate* con *night rate* o *hat rack*,¹⁴ que no tienen cúmulo). Esta distinta manera de articular se aprecia bien en *sparkling* frente a *starkly*: la primera se deriva de una CS y tiene inicio ramificado (spar·kling), mientras que la segunda no tiene CS de base, y el derivado no tiene inicio ramificado (stark·ly).

La clave para formar el inicio ramificado está en la estructura¹⁵. Es lo que tienen en común *mystery* y *possibly* frente a *sharply* (al margen del sufijo, por tanto). Hay que empezar por la **base**, antes de que se agregue la vocal por la derecha. Vemos entonces que *sharp* termina en una coda sin más, en tanto que *possible* tiene una CS, y *mystery*

admite tanto schwa como CS¹⁶. Recordemos que las CS representan el paso intermedio de todo un espectro de reducción: están a caballo entre la schwa y la supresión. Y esta es la solución.

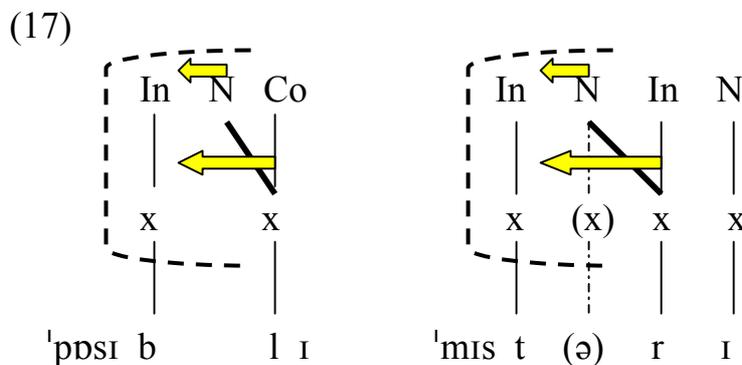
La CS se asocia al N previo, y este está en la misma sílaba que el inicio que se va ramificar:



En este ámbito homosilábico es donde el inicio puede ramificar. Nos encontramos en el **plano silábico** y aquí entra en acción el **Principio de Maximización de Inicios**, como vimos en el capítulo anterior. El inicio ramificado se forma de manera natural cuando la reducción alcanza el nivel de la supresión. Porque de esto se trata: es un caso particular de la supresión, y sólo requiere de unas condiciones fonotácticas favorables (como en *amply*, *singlet*, *smuggling*, *tackling*, *stapler*, *simplest*). Estas condiciones son también las que bloquean el inicio ramificado en *camera*, *javelin*, *company*, o *decimal* y muchas otras (ver Harris 1994: 185, diagrama 50.b), pero es simplemente porque en inglés no existen los inicios /mr-/ , /vl-/ , /pn-/ y /sm-/ , respectivamente¹⁷.

Volviendo a la estructura, no dejaremos de insistir en la necesidad de considerar el proceso en toda su extensión. Hay una reducción

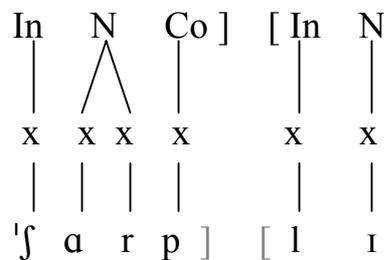
progresiva, que pasa de la vocal plena (V) a la schwa (v), después a la CS y finalmente a la supresión (\emptyset). La estructura de nodos se hace eco de esta condensación progresiva, y se va replegando sobre sí misma (es muy apreciable la pérdida del aspa autosegmental en la CS) hasta que finalmente se pierde una sílaba (en la supresión). En la formación de un inicio ramificado, la supresión produce un repliegue que remeda la **afijación telescópica** (*affix telescoping*, ver Haspelmath (2002: 56)), de modo que esta condensación silábica podría llamarse ‘**silabeo telescópico**’. Su plasmación gráfica es:



Como venimos defendiendo (recuérdese el análisis que dedicamos a la coalescencia), lo primero que se pierde es el material melódico, y al no tener nada ligado se pierde también la estructura superior. Pero no al contrario. En lo que respecta a los nodos, el repliegue toma como pivote el primer inicio, que será el que encabece el nodo ramificado una vez que este se forme. Recuperamos así la idea de un nodo que actúa como **atractor**; esta vez se trata de un nodo inicio, que en este caso queda a la izquierda de la CS. En conjunto es un fenómeno estructural (la cabeza está a la izquierda), aunque implica factores puramente fónicos, en relación con la composición de los segmentos afectados.

Pero al igual que las condiciones fonotácticas desfavorables impedían la formación de un inicio ramificado, existen también casos de condiciones favorables donde no se forma. Esto es achacable a la estructura. Cuando no hay un núcleo que pueda replegarse y ser absorbido por el inicio de la izquierda, las consonantes no forman cúmulo. El Principio de Maximización no puede actuar y no hay inicio ramificado; por eso en *sharply* o *starkly* hay simplemente una secuencia coda]·[inicio, en sílabas diferentes:

(18)



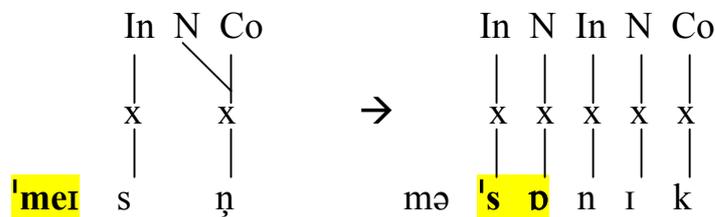
Y diremos de pasada que hay casos en que no se produce condensación silábica, aun cumpliéndose las condiciones silábicas y fonotácticas necesarias. Por ejemplo, *hickory* o *chicory* conservan la schwa, hecho que nosotros atribuimos a factores cognitivos, y más concretamente al reconocimiento, que será el tema de la siguiente sección.

Una consideración final. El proceso de reducción progresiva, llevado hasta la condensación de nodos, debe de funcionar como mecanismo de **adaptación evolutiva**, de forma similar a los que tienen los seres vivos. Entendemos que el aparato fonológico tiene que ser capaz de hacer frente a circunstancias lingüísticas y extralingüísticas que impongan ajustes en las unidades léxicas y en las estructuras y reglas

del sistema. Como destaca Aitchison (1991), las lenguas no son entes estáticos, sino que están en permanente evolución (es consustancial a ellas), y a tal efecto el aparato (sincrónico) tiene que estar preparado para hacer los cambios (por ejemplo, incorporar préstamos, tanto léxicos como gramaticales) igual que los seres vivos se adaptan al entorno. Y tanto en estos como en el lenguaje, no se concibe la evolución sin una maquinaria estructural sincrónica que maneje los cambios pertinentes¹⁸.

CUARTA OPCIÓN: La CS alterna con una sílaba tónica, que tiene la vocal plena. Es lo que sucede en el paso de *Essence* a *eSSEntial*, o de *MYStery* a *mysTERious*, así como en el sufijo {-ic}, que desplaza el acento a la sílaba previa (p. ej., al pasar de *MAson* a *maSONic*)¹⁹:

(19)



El núcleo al que estaba vinculada la CS recibe el acento principal de palabra. En estas condiciones, la vocal se expresa plenamente, y la CS no se puede vincular a ese núcleo²⁰. Esta alternancia entre CS y vocal tónica, plenamente especificada, es un buen indicador de que en la parte baja de la estructura siempre está la vocal subyacente²¹.

Una vez expuestas las cuatro opciones del contexto derecho, nuestro análisis va a completarse con el papel que juega la cognición a la hora

de elegir una variante u otra; veremos que no todo puede imputarse a la estructura.

Condiciones morfofonológicas: reconocimiento

Si el contexto derecho ya era de por sí complicado, por la existencia de cuatro opciones de realización, hay dos factores adicionales que vienen a complicarlo un poco más. Vamos a examinar unos cuantos ejemplos concretos en una tabla comparativa, para ver en qué consiste la dificultad:

(20)

BASE + (SUFIJOS)	SCHWA	CS	SUPRESIÓN
Fission(able)	[-ʃənəb̩]	-ʃnəb̩	—
Excursion(ist)(ize)	[-ʃənɪst], [-ʃənaɪz]	-ʃnɪst, -ʃnaɪz	—
Reception(ist)	-ʃənɪst	[-ʃnɪst]	[-ʃ·nɪst]
Reason(able)(ing)	— [-zənɪŋ]	-znəb̩l, -znɪŋ	[-z·nəb̩], [-z nɪŋ]
Orphan(age)	ˈɔːfənɪdʒ	[-fnɪdʒ]	—
Pardon(ing)(er)(able)	—	-dnɪŋ, -dnə, -dnəb̩	-dn·ɪŋ, —, -dn·əb̩

En la primera columna hemos colocado las **formas básicas**, a las cuales se les añadirá un **sufijo flexivo o derivativo** (indicado entre paréntesis) que aporta un contexto vocálico. Todas estas formas

básicas son netamente silábicas, tienen una CS; por ejemplo, *fission* es [ˈfɪʃən].

Las otras tres columnas muestran las formas derivadas posibles para cada base. Vemos que todas ellas tienen una variante con CS, y lo que marca la diferencia es con quién alterna esa CS: en unos casos alterna con schwa, en otros con supresión y en otros con schwa y supresión. Un segundo factor de complicación es que hay palabras que muestran preferencia (marcada en negrita), por una de las variantes de la oscilación, en tanto que otras palabras prefieren la otra variante. [La(s) forma(s) secundaria(s) van entre corchetes.]

Es evidente que la alternancia de formas no obedece al contexto, a diferencia de las alternancias clásicas de la fonología. Y ya que todo ocurre en un único contexto, descartamos que la variación dependa de las condiciones fonotácticas. ¿Por qué se elige entonces una u otra posibilidad o bien por qué se permite cierta alternancia en unos casos pero en otros es inviable, cuando las condiciones parecen idénticas para todos?

El sufijo no parece que desempeñe un papel importante, porque un mismo sufijo produce patrones diversos. Así, la CS es la variante principal para *excursionist*, mientras que en *receptionist* la variante principal es con schwa; además, *receptionist* tiene también supresión, en tanto que *excursionist* no la admite. También vemos cómo sufijos distintos producen un mismo patrón (con la base *excursion*), y no afecta demasiado que se trate de flexión o derivación, a juzgar por la base *reason*.

Pasamos a la base. Reparemos en el hecho de que la supresión es la opción menos favorecida. Un problema de esta opción es que se puede producir una forma que coincide con otra que ya existe, por lo cual se **bloquea** el resultado; por ejemplo, *finally* se podría confundir con *finely* tras la supresión²². Dando un alcance más amplio al criterio, podemos suponer que *formally* no tiene supresión para no confundirse con *formly*, una palabra que no existe pero que es una palabra posible del inglés. Esta confusión producida por una palabra inexistente apunta a la **raíz morfológica** como clave de la supresión.

Comparando más ejemplos, vemos que la variante principal de *stiffening* y *cushioning* es una CS, pero la variante secundaria para *stiffening* es una supresión ['stɪf·nɪŋ], en tanto que *cushioning* no admite supresión y necesita una schwa, ['kʊʃənɪŋ]. No se pueden aducir razones estructurales (de legitimación o de formación de núcleos vacíos) para justificar esta diferencia, porque *cushioning* y *stiffening* son análogas.

La clave de la supresión debe de estar en la **morfología**. En *stiffen* hay una raíz *stiff-* independiente, mientras que *cushion* es monomorfémica y no puede aislarse *cush-*. Cuando elidimos hasta llegar a la supresión, es sencillo reconocer *stiff* en ['stɪf·nɪŋ], pero no hay nada que reconocer en ['kʊʃ·nɪŋ], y por eso no hay supresión.

Aun así, no es tan simple, porque *fashionable* sí que admite la supresión y *fashioning* permite las tres opciones, a pesar de que *fashion* es tan monomorfémica como *cushion*. ¿Cómo es posible? Nuestra intuición es que ['fæʃ-] nos remite a *fashion*, a pesar de no ser

una raíz o un componente aislable de esa palabra. Al oír [ˈfæʃ-] pensamos fácilmente en *fashion*. No ocurre lo mismo con [ˈkʊʃ-], que no sugiere *cushion* [ˈkʊʃŋ]. Podemos **reconstruir** *fashion* a partir de *fash-*, pero no podemos reconstruir *cushion* a partir de *cush-*, y por eso hay que expresar una schwa en *cushioning*.

Sorprendentemente, hay palabras en las que la schwa no facilita el reconocimiento, sino todo lo contrario. Veamos un caso. Los derivados de la base *pardon* admiten CS (la variante principal es /-dŋ/) y pueden tener supresión²³, pero ninguno de los derivados tiene la forma con schwa, y es que puede ser que sea más fácil reconocer *pard-* que *pardə-* como la clave que remite a *pardon*, curiosamente. Por ende, la schwa no siempre es una opción disponible, aunque parezca que debería serlo. Tampoco es posible la schwa en *reasonable*, como vemos en la tabla; será que [ˈri:z-] es más identificable con *reason* que [ˈri:zə-].

La conclusión es que a la hora de favorecer una reducción fónica, debe de jugar un papel importante la probabilidad de reconocer o reconstruir la palabra original (esto es, la forma sin reducir). Si la forma que se reconoce o reconstruye coincide con otra forma que ya existe, se produce el bloqueo: en *fissionable* se bloquea la supresión porque remite a *fish*.²⁴ Y si la forma que se usa como base para el reconocimiento no sugiere nada, también se bloquea la supresión. Por ejemplo, no se permite supresión en *mentally* porque *ment-* no dice nada, otra vez cuesta reconstruir la forma original, por mucho que /ˈment·li/ esté bien formado fonológicamente. Podría trasladarse a la morfofonología una versión del **Principio de Percepción** y

Conservación del Contorno que Jackendoff (1993: 178-9) sugiere para el procesamiento cerebral de la visión. Esta manera de interpretar la morfología es novedosa, hasta donde sabemos, porque tiene un componente cognitivo y no se limita a las raíces gramaticales del diccionario. Aunque el reconocimiento no puede explicarlo todo, al menos ayuda a entender las preferencias de la tabla, y creemos que también es admisible descartar la influencia de las condiciones fonotácticas en ese caso²⁵.

Hay una relación importante entre estructura y cognición. Porque el reconocimiento lo que hace es establecer una conexión entre la forma percibida y la forma de base, subyacente (esto se aplica a *bottling*, por ejemplo). Y por otra parte, el mejor alineamiento no se limita a la estructura, porque al favorecer el procesamiento interesa al componente cognitivo. Estas dos fuerzas (reconocimiento y alineamiento) ponen los aspectos estructurales al servicio de la cognición cerebral, y la pugna entre ambas obedece a este criterio.

Nuestra propuesta está asumiendo además una nueva forma de entender la fonología, que se añade a lo que se ha hecho hasta ahora. Igual que existe un ‘camino de ida y vuelta’ para la fonética, que tiene un componente dedicado a la producción de sonidos (**articulatoria**) y otro a su percepción (**perceptiva**), lo mismo debería aplicarse a la fonología. Los estudios en este campo se centran siempre en la **producción fonológica**, desde los elementos más básicos hasta las estructuras más complejas. Pero creemos que también hay que prestar atención al proceso inverso, aquel que se ocupa de **reconocer formas** y **asignarles representaciones** fonológicas en nuestro cerebro o mente. Podríamos así hablar de **fonología productiva** y **fonología**

receptiva, o bien fonología **aferente** y **eferente**. Y acaso pueda asumirse el mismo planteamiento para la sintaxis.

Hay que matizar que la fonología receptiva no se ocupa de reconocer sonidos, tarea esta que compete a la fonética acústica. Lo que hay que reconocer son conceptos (ítems) mentales. Se trata de acoplar (*match*) la representación mental de una aferencia auditiva con un concepto que tengamos ya almacenado en la memoria cerebral, con una representación fonológica o morfosintáctica. Nosotros atribuimos a esta fonología receptiva la formación de la **etimología popular** (*folk etymology*), una manera de formar palabras que depende de la asignación de una secuencia fónica a un concepto ya conocido. Como señala Algeo (2010: 241), este proceso es bien frecuente en niños, y cuando las formas cristalizan y se hacen de uso común, es cuando hablamos de etimología popular. La clave está en el reconocimiento y la asignación de formas, que no es una simple cuestión acústica, sino cognitiva.

¹ Roca (1994: 134) cita este patrón como el principal argumento a favor de la existencia de las sílabas: *The main argument for the postulation of syllables hinges on the recurrence of the context {#, C} in rules.*

² (1995: 39): *the bottom row of x's (which serve only as place markers) is eliminated.*

³ Hay una circunstancia en la que el contexto consonántico /_ C favorece aún más la aparición de una CS. Vimos cómo en las antigeminoides la CS nasal es fuertemente silábica tras asimilarse al punto de articulación del inicio que está a su izquierda. Pues bien, cuando el punto de articulación de la izquierda coincide además con el de la derecha, la asimilación de LUGAR no sólo es posible, sino prácticamente inevitable. Por ejemplo, la /ŋ/ de *davenport* se encuentra entre dos consonantes labiales, y por ello tiende a asimilar este punto de articulación. Igualmente, en *bacon crisps*, a través de la linde de palabra, puede favorecerse la articulación velar.

⁴ Recordemos el ejemplo de *Human Rights Watch annual conference* que citamos en el capítulo 8.

⁵ Por cierto que la asignación silábica no es igual cuando la /__V está en una palabra distinta, separada. Por ejemplo, en *little ants* o *mountain oaks* se conserva siempre la silabicidad de la CS. La CS está en su coda de origen, y no se maximiza el inicio, por lo que la silabicidad no se afecta. Así como las **lindes internas** (de morfema) no influían en la estructura, las **lindes externas** (de palabra) sí que lo hacen.

⁶ Harris ya apunta en esta dirección. Pero no traza la estructura de pies (que sólo queda esbozada) y lo explica todo en términos de **núcleos vacíos** y **legitimación** (*licensing*) entre núcleos, más pegada al nivel segmental. Nuestro modelo integrado tiene, a nuestro entender, mejor capacidad de explicación.

⁷ Aquí nuestro análisis se aparta, de manera fundamental, del de Harris (1994: 183), que considera que el adjetivo *separate* está en un pie ternario (con estructura plana), al que denomina **superpié**.

⁸ Nótese que la supresión, con la eliminación del N de la CS, sólo es posible en /__V. Porque al ser ambisilábica, la R tiene otro N al que asociarse. Sin embargo, en una CS/__C, la R quedaría “huérfana” en su coda.

⁹ Harris lo define en términos de núcleo vacío interpuesto, y también incluye ejemplos en los que el inicio no puede ramificar, como en *badly*. Nosotros preferimos restringirnos en este momento a los casos en que la coda es menos sonora que el inicio que le sigue, como en *chutney* (1994: 223) o en los ejemplos ya citados.

¹⁰ Los datos de los diccionarios apuntan además a una cierta flexibilidad en las variantes. De nuevo tenemos la impresión de que no es ‘todo o nada’, sino que hay grados o porcentajes de realización. Por ejemplo, el diccionario *Oxford* da *open* y *openly* como no-silábicas, pero *opener*, *opening* tienen supresión, y *openness* es silábica. Y para el *Webster*, *open(ed)* admite la silábica y la no-silábica, pero *openness* y *openly* son sólo no-silábicas, y para *opener*, *opening* hay tanto no-silábicas como supresión.

Si alguna de las formas de la oscilación es más inestable que otras, o simplemente se prefiere a las otras por algún motivo, como recoge el diccionario *Everyman’s*, estaremos también ante un proceso de evolución histórica. Las diferencias léxicas y las formas ligeramente anómalas o no-del-todo-bien-formadas pueden ser un motor del cambio diacrónico, igual que en la evolución de las especies biológicas

¹¹ Un detalle de estas triples alternancias es que la CS parece tener un silabeo distinto del de las otras dos formas (la suprimida y la de la schwa). Por ejemplo, compárese la forma con CS [tʃɑ:n.s]ə], donde la última sílaba no tiene inicio, con [tʃɑ:n.sə.lə] y [tʃɑ:ns.lə], que tienen los inicios maximizados. Da la impresión de que la CS se ha desvinculado de su inicio de origen y ocupa el núcleo previo, hecho que parece avalar la tesis de la supresión con resilabeo.

¹² Históricamente sí deriva de un inicio ramificado, pero tuvo que pasar a [-bəl] en el curso del tiempo. Si actualmente estuviera en un inicio ramificado, sería una **consonante atrapada** según el planteamiento de Scheer (2003b).

¹³ La nota 59 del capítulo 4 de Harris (1994) hace referencia a esta dificultad.

¹⁴ La formación del cúmulo es sensible a la linde de palabra, pero no a la linde silábica. Tampoco puede interponerse una schwa o un núcleo vacío, porque el cúmulo requiere adyacencia estricta. Análogamente, en la sintaxis oracional no se puede contraer un auxiliar negativo cuando se interpone una categoría vacía.

¹⁵ Harris (1994: 192) menciona pares de palabras cuyo primer miembro tiene schwa y el segundo tiene un inicio ramificado (*couple/couplet*, o *meter/metric*, *simple/simplicity*). Y hace notar, con razón, que en ninguno de estos ejemplos se puede expresar una schwa; por ejemplo, **coup[ə]let*, **met[ə]ric*. También cita un análisis por el que en *couple* hay un cúmulo final que se rompe al entrar una schwa epentética, y así $k\lambda pl \rightarrow k\lambda p\lambda l$, por lo tanto sin CS en ningún caso.

Sostiene Harris que estas alternancias deben ir *simply listed in the lexicon*, porque las formas con inicio ramificado son fonológicamente indistinguibles de palabras como *poplar*, *petrol*, *patron*, que no tienen alternancias radicales. Si la lista fuera tan corta y poco productiva, la postura estaría bien justificada, pero la flexión con el gerundio {-ing} por sí sola produce un número demasiado amplio de nuevas formas (p.ej. *sparkling* o *stapling*) “fonológicamente indistinguibles” de *poplar* o *petrol*, a las que hay que sumar las que se producen por derivación morfológica en todos los casos de {ble}+{ly} → {bly} y en otras derivaciones (p. ej. *stapler*), y que sería excesivo especificar de una en una en el diccionario.

¹⁶ Los ejemplos con /ɹ/ muestran un proceso irreversible, pero sin embargo con la /ɹ/ hay oscilación resonante, como sucede en *mystery*, *natural*, *interesting*, o *factory*

Hay un número razonable (y va en aumento) de casos de síncope que se va congelando léxicamente y va dejando de alternar con otras formas. Por ejemplo, en la palabra casi exclusivamente británica *pram* (<*perambulator*>), la única forma posible es /præm/. La palabra *every* (<*everich*>) ya es casi siempre es ['evrɪ], aunque en ‘*Losing my religion*’ de REM se oye sucesivamente ['evərɪ] y ['evrɪ], cuando dice “*Every whisper, every waking hour (...)*”. Para *several*, el *Everyman’s* da sólo la forma sin schwa, mientras que el *Webster* acepta las dos variantes, con y sin schwa. *Interest(ing)* tiende a no tener schwa. En portugués brasileño, *para* es casi siempre /pra/.

Harris (1994: 192) habla del grado de obligatoriedad de síncope (y CS) en distintas lenguas. Por ejemplo, las síncope son obligatorias en tonkawa (una lengua amerindia de Texas), mientras que en español es imposible formar CS. En inglés son opcionales tanto las síncope como las CS, según llevamos visto.

En inglés la opcionalidad tiene a veces forma de lo que en genética se denomina **codominancia**: encontramos una forma o la otra, pero nunca las dos juntas, en oscilación. Es así, por ejemplo, en *Barbara* (Stanwyck) frente a *Barbra* (Streisand) o en *Deborah* (Kerr) frente a *Debra* (Winger). También en checo y eslovaco existe el doblete *Peter* y *Petr*.

¹⁷ Harris no da espacio para relajar las reglas, pero es muy probable que se tolere un pequeño grado de desviación, como hace notar Jackendoff (1993: 183) cuando dice que vivimos en un mundo que no cuadra exactamente en nuestros patrones mentales, y por lo tanto debemos ser capaces de hacer algo ante las pequeñas desviaciones. Un ejemplo de lo que estamos diciendo es la pronunciación mencionada de *every*, que a pesar del inicio atípico /vr-/ ya está ya muy fijada como ['evrɪ], especialmente en los derivados como *everything* o *everybody*; y no sería extraño que *several* siguiera el mismo camino.

¹⁸ De nuevo la idea de **evolución desde dentro** que propuso Stephen Jay Gould, y que ya explicamos en las notas del capítulo 8.

¹⁹ Aunque la opción con sílaba tónica no se presta a la triple alternancia, sí que la vemos en *MAMmal*, *mamMALian*, y *mammaLIferous*, cuyas variantes tienen CS, vocal tónica y schwa, respectivamente.

Otro detalle en relación con la alternancia entre CS y vocal tónica es que no es necesario que se añada un contexto vocálico para que ocurra, como vemos en el paso de nombre a verbo que hay en [PREsent]_N → [preSENT]_V. En realidad, la alternancia entre CS y vocal tónica es posible incluso dentro de una misma palabra, sin que haya derivación morfológica, como demuestra la doble acentuación (y el doble silabeo) que se propone en Webster para el contraste entre /bicenTE · nary/ y /biCENTɹ · ary/.

²⁰ También cambia la estructura métrica, porque el troqueo de (MA son) se desplaza hacia (ma) (SO nic).

²¹ La alternancia entre una CS y una vocal plenamente especificada es un argumento a favor de la reducción frente a la epéntesis, como vimos en el capítulo 2. El motivo es que no puede predecirse cuál va a ser la vocal plena que reemplace a la CS.

Aún se puede postular que, aunque la vocal plena se especifique léxicamente y tenga que aprenderse, en la mente de los hablantes no se pasa de una forma subyacente a otra derivada. Por ejemplo, si preguntamos a hablantes analfabetos (con objeto de que no puedan recurrir a la ortografía para sacar conclusiones) cuál sería el resultado de añadir el sufijo {-ic} a la base *nation*, creemos que lo más probable es que tuvieran que inventarse la forma “subyacente”, de modo que el resultado variaría de unos hablantes a otros.

²² También es sabido que en el gerundio *coddling* (de *coddle* ‘hacer mimos’) la /l/ debe seguir siendo silábica para que no se confunda con el nombre *codling* ‘bacalao pequeño’.

²³ Admiten supresión *-ning* y *-nable*, aunque no así *-ner*, quizá porque *-ing* y *-ble* sean más reconocibles como sufijos que *-er*, que al fin y al cabo queda como una simple [-ə] en los dialectos no-róticos.

²⁴ Además, *fissionable* es una palabra técnica y poco frecuente, al contrario que *fish*, con lo que está en desventaja como candidata al reconocimiento.

²⁵ Ramachandran (2011: 57-60) dice que *the exact manner in which object recognition is achieved is still quite mysterious*. Cita como posibles factores las relaciones, las plantillas (con patrones o modelos) y la exageración. También la información almacenada en la memoria del cerebro. Pero al fin y al cabo, *the word “recognition”, however, doesn’t tell us anything much unless we can explain how the object or the face in question evokes meaning –based on the memory associations of the face. The question of how neurons encode meaning and evoke all the semantic associations of an object is the holy grail of neuroscience, whether you are studying memory, perception, art, or consciousness.*



11 Las CS menos típicas

Las otras CS del inglés

Hay en inglés un tipo de CS que nos resultan peculiares, y no sabemos de nadie que las haya considerado así hasta ahora. Se trata de la /m/ silábica de *rhythm* y de toda la pléyade de /-sm/ que podemos ilustrar con *chasm*, *prism*, *citoplasm*, *microcosm* o *minimalism*. Estas CS parecen como las demás, con su alternancia entre schwa y silábica, pero tienen características que las distinguen. Llama la atención que su posición sea exclusivamente final, y el motivo es que no toleran la presencia de una V a su derecha. Cuando por derivación morfológica se añade un contexto vocálico (en *rhythmic* o *schismatic*, por ejemplo), se pierde al instante la silabicidad, y la /m/ queda como inicio de la vocal añadida.

Más significativo aún es que el contexto vocálico no sólo impida la silabicidad, sino también la schwa. Esta vocal se forma por reducción de una vocal plena, como sabemos, pero en las CS atípicas no se atisba cuál pueda ser esa vocal plena “de origen”. Al añadir un sufijo atractor del acento, como es #-ic#, la vocal plena debería salir a la superficie, como en *system* → *systEmic*. Pero en *rhythm* → **rhythEmic* no sólo no hay vocal plena, sino que no hay vocal en absoluto, ni siquiera la schwa. Incluso podemos comparar cómo la

schwa se mantiene en *problem* → *problematic* pero desaparece por completo en *schism* → *schis-matic*. Si la schwa de las CS atípicas no se ha formado por reducción de una vocal plena, ¿de dónde proviene entonces? Con la perspectiva que nos da todo el análisis que llevamos realizado, vamos a estudiar estas **CS atípicas** más de cerca, y obtendremos conclusiones sorprendentes.

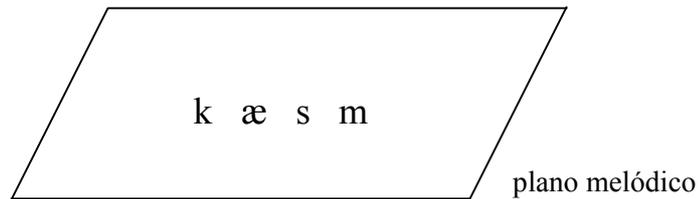
Hemos visto cómo las **CS típicas** del inglés se forman en virtud de un mecanismo métrico que alterna partes de ‘luz’ y de ‘sombra’, con el fin de facilitar el procesamiento cognitivo de las secuencias fónicas. En este contexto, las CS se forman en el componente métrico de la derivación fonológica, en el plano más alto. Porque es aquí donde se forman pies trocaicos, en cuya parte débil la vocal se reduce a schwa y a su vez puede coalescer con la consonante de la coda para formar una CS.

Pero hemos visto que en las CS atípicas la schwa no proviene de una reducción vocálica. Así que la motivación para formarlas no es métrica, y por tanto no tienen que ver con el patrón de luz-sombra de procesamiento cognitivo. Concluimos entonces que estamos ante un nuevo porqué en la formación de CS inglesas, y vamos a ver cómo esta nueva motivación se remonta en último término al componente cognitivo de la economía mental.

Las CS del tipo /-sm/ provienen de préstamos grecolatinos donde se ha eliminado material melódico. En origen estas palabras terminaban en vocal, pero al suprimirla se queda la /-m/ como final de palabra. Con esta secuencia de segmentos se consigna la palabra en el lexicón,

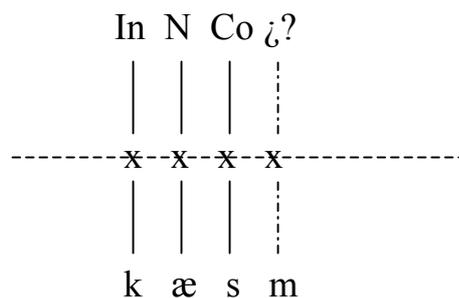
y como tal debe figurar en el componente melódico al inicio de la derivación fonológica; por ejemplo, para *chasm* tenemos:

(1)



La secuencia debe someterse a las reglas de silabeo ordinarias... Pero de inmediato surge un conflicto con la consonante final, que rompe el perfil de secuenciación sonora necesario para licenciar una melodía:

(2)



Y es de toda lógica que se introduzca entonces un mecanismo reparador para que se pueda **legitimar** la secuencia¹. O sea que estamos ante un caso que parece encajar en la explicación que ha venido dando la Fonología de la Rección (FR) sobre la formación de CS.

Pero las cosas no son tan claras. Las secuencias conflictivas de [obstruyente + m] también se producen en otras palabras, y sin embargo no se recurre a formar una CS para resolver el problema. Estamos pensando en *phlegm* o en *diaphragm*, /flem/ y /^ldaɪəfræm/,

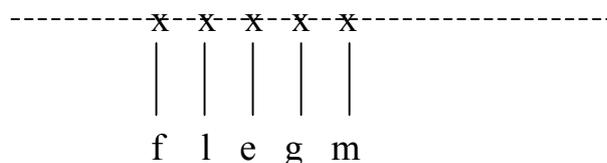
que recurren a eliminar la /g/ subyacente para evitar el colapso sonoro. El indicio o prueba de que estas palabras mantienen una /g/ subyacente (y no puramente ortográfica, como podría aducirse) es que sus derivados morfológicos *phlegmatic* o *diaphragmatic* mantienen plenamente la /g/ cuando deja de haber conflicto sonoro. Es obligado preguntarse por qué se da un tratamiento distinto (*phlegm* respecto a *chasm*) para un mismo problema.

Aún más todavía. Las propias CS de *chasm* y *rhythm* tienen que tener otra justificación que la de la reparación fonológica. Porque todas las palabras de este tipo tienen una forma con schwa (/ˈkæəsəm/ y /ˈrɪðəm/) que bastaría por sí sola para enmendar el error. Si a pesar de todo se forma una CS, además de la forma con schwa, ¿por qué o para qué se hace?

Nuestra idea es que el problema de base está bien definido: hay una secuencia final [obstruyente + m] que no permite cumplir con el **Principio de Secuenciación Sonora (PSS)**, y que impide la legitimación de todos los segmentos mediante el silabeo. Estamos en la parte más baja de la estructura, en la transición entre el plano melódico y el plano silábico.

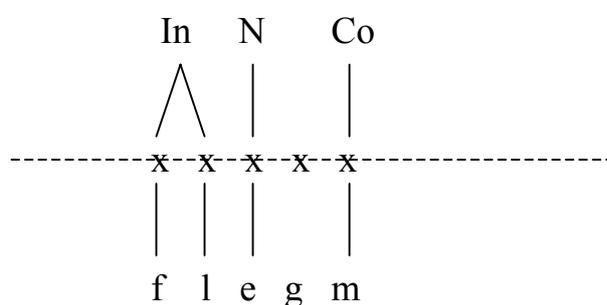
Para mayor facilidad de exposición vamos a centrarnos en *phlegm* antes de examinar las CS del tipo *chasm*. En el plano melódico tenemos una secuencia de segmentos [f-l-e-g-m] que viene dada léxicamente:

(3)



donde los dos segmentos finales plantean un problema de silabeo, como hemos visto. Una solución es eliminar uno de ellos, y en una primera ojeada es lo que se hace. Eso sí, la eliminación no es total, puesto que el segmento que no aparece realizado en /flem/ sí que lo hace en /fleg¹mætik²/. ¿Pudiera ser que se deje como segmento flotante dentro de la estructura, en vez de eliminarlo?

(4)



El diagrama es muy claro, pero tiene algún inconveniente. Por sí solo no explica por qué se omite la /g/ y no la /m/, que también resolvería el problema. Es más, todas las secuencias problemáticas /-gm, -ðm, -sm/ tienen en común el segmento /m/; si lo eliminamos de todas ellas, se acabó el problema. Pero no es lo que se observa, y el motivo es que se resolvería un problema a costa de crear otro mayor, según nuestro criterio.

El problema que se resuelve es fonológico, el de la secuenciación sonora y el silabeo. Pero el problema que se crearía es de orden cognitivo: está en juego el **reconocimiento** de las formas básicas, subyacentes. Porque /fleg/, /rɪð/, /kæs/ no remiten a los originales *phlegm*, *rhythm*, *chasm*. Esta es la clave. Y no es sólo por la escasez de segmentos constituyentes, que apenas puede permitirse el lujo de eliminar material sin resentirse. También en palabras más largas, donde sería más fácil rastrear el original, como en *totalitarianism*, no es posible suprimir la /m/ final sin tener la impresión de que nos falta algo, y aún se prestaría a confusión con su par mínimo *totalitarianist*.

En definitiva, no podemos eliminar alegremente la /m/ final. Desde el punto de vista estrictamente fonológico, no habría mayor problema; el diagrama con el segmento flotante permitiría realizar esta operación y otras más. Por eso el componente fonológico muestra los mecanismos que son posibles, pero no explica por qué se opta por una regla y no por otra. Para dar una explicación hay que introducir un factor cognitivo, que es el que motiva las operaciones fonológicas realizadas.

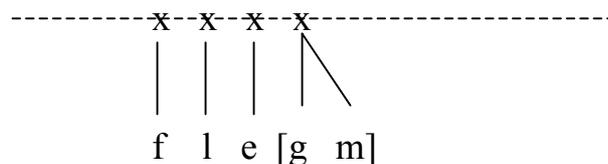
Volviendo al diagrama, la solución del segmento flotante tiene otro inconveniente añadido, a pesar de su atractivo inicial. ¿Por qué no se aplica a todas las secuencias problemáticas [-gm, -ðm, -sm]? De nuevo parece que hay un problema de reconocimiento. Así como /flem/ remite muy bien a /flegm/ *phlegm*, no ocurre igual con /rim/ respecto a /rɪðm/ *rhythm* (que además se **bloquea** con /rim/ *rim*) ni con /kæm/ respecto a /kæsm/ *chasm*. Otra vez da la sensación de que se pierde integridad fonológica y se dificulta el reconocimiento muy

gravemente. Aún más claro lo vemos al invertir la pregunta: ¿por qué /flem/ remite a /flegm/ *phlegm* sin mayor problema?

La respuesta es fonológica esta vez, aunque en última instancia tiene consecuencias para la cognición. Los dos segmentos finales de /flegm/ son muy parecidos entre sí. Hasta el punto de que si en vez de una /g/ tuviéramos una /b/, sería una geminada parcial. Aún más, /g/ y /b/ difieren sólo en el lugar de articulación, pero incluso en esto comparten el ser no-coronales. Dicho esto, podemos concluir que los dos segmentos finales de /flegm/ son muy parecidos entre sí, algo que no ocurre con /riðm/ y menos aún con /kæsm/, que difieren en más rasgos todavía. ¿En qué afecta este dato a la solución que se adopta?

Descartamos la hipótesis del segmento flotante, y en su lugar proponemos que en /flegm/ lo que se produce es una **coalescencia** de segmentos. Dos segmentos muy análogos confluyen en uno solo, en una suerte de **haplología** de geminadas:

(5)



Intuimos que la razón de que confluyan en la /m/ y no en la /g/ es otra vez cognitiva, por la facilidad para reconocer la forma subyacente, aunque puede haber factores fonológicos implicados también³. Pero dado que en *autumn* se opta por absorber la /n/, no parece que la posición en la estructura melódica o silábica sea el factor

determinante. Aun así, quedamos a la espera de nuevos avances en la investigación para arrojar más luz sobre este aspecto.

Mecanismos de formación de las CS atípicas

Una vez resuelto el problema de secuenciación sonora en *phlegm*, hace falta encontrar una solución para los casos que no pudieron acogerse a este remedio. El hecho es que hay una schwa en /'rɪðəm/ y /'kæsəm/ (para *rhythm* y *chasm* respectivamente), y si no proviene de una reducción vocálica, tendremos que asumir que esa schwa se insertó en la secuencia de base. De inmediato vuelve la pregunta de por qué no se adoptó esta solución como forma universal, abarcando también a *phlegm*⁴.

La respuesta debe de ser de economía del sistema fonológico: la **epéntesis** o inserción de un segmento parece una solución más “**cara**” que la coalescencia de dos segmentos análogos. Como indicio de que esto es cierto, podemos traer a la memoria varios casos de coalescencia en inglés, pero rara vez de inserción de vocales. Así que en *phlegm* se optó por la solución más económica o más “**barata**”, y en *rhythm* y *chasm* hubo que recurrir al siguiente paso, más “caro”, al no bastar con lo más “barato”.

Hay además un aspecto cognitivo que no podemos desdeñar. Pensamos que las CS atípicas aprovechan el mecanismo de las CS, que el sistema ya tenía. Digamos pues que las CS atípicas (y sus variantes con schwa) se forman por **analogía** con otras formas (y reglas de formación) que ya existen en inglés y que les sirven como

fuentes de inspiración. Y una vez metidos en la analogía, nos parece que *chasm* y *rhythm* emulan el /-səm/ y el /-ðəm/ que ya existen en *blossom* o *fathom*. Por el contrario, no se ve cuál podría ser el modelo de /-gəm/ para que *phlegm* diera /'flegəm/, y esto quizá influya para formar unas secuencias y no formar otras.

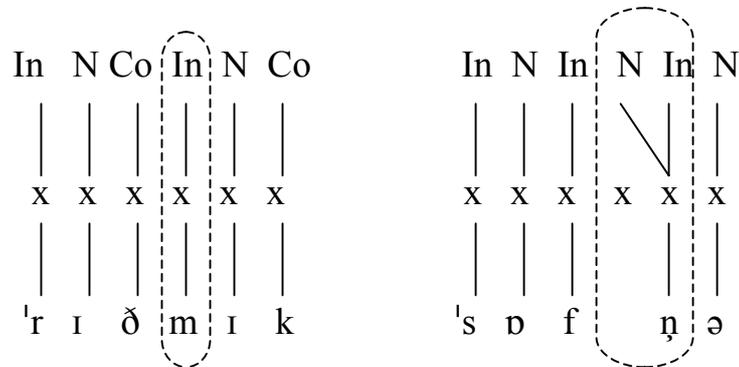
Suponiendo, pues, que haya una schwa epentética en *rhythm* y *chasm*, hay que examinar cómo se produce ⁵. Podríamos estipular la inserción lisa y llana de un segmento en la secuencia sonora, pero un mecanismo así sería demasiado potente. Hace falta establecer restricciones para evitar el ‘todo vale’. La schwa está ahí, porque es perfectamente audible en /'rɪðəm/ o /'kæsəm/, pero no nos conformamos con una inserción directa del segmento. Hay que analizar la estructura.

Una pista la encontramos en los derivados morfológicos de *chasm* y *rhythm* que tienen vocal y ya no presentan problema con el contorno sonoro. Por ejemplo, en *rhythmic* o *schismatic* se hace un silabeo normal. Es notorio que la /m/ ya no es silábica en estos derivados (y aún más: es que no puede seguir siendo silábica), en contraste con las CS habituales, que son capaces de mantener la silabicidad en contexto vocálico /__V, como en *softener* /'sɒfɪnə/. Alguna diferencia de estructura tiene que haber, entonces.

Si las formas con /__V del tipo *rhythmic* o *schismatic* tienen un silabeo normal, y es imposible que la /m/ sea silábica en ningún caso, hay que suponer que está asociada exclusivamente a un nodo inicio.

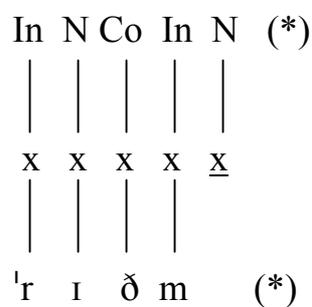
En contraste, las CS habituales son ambisilábicas y se asocian a la vez a un inicio y al núcleo previo:

(6)



La /m/ de *rhythmic* sólo tiene asociación al nodo In, y funciona en relación con el núcleo siguiente pero no con el precedente. Cabe pensar que hay algo de esto en las formas conflictivas, como *rhythm*: que esa /-m/ final debería ser en realidad el inicio del núcleo siguiente, pero se encuentra con que está vacío. El argumento casa bien con la supresión de la vocal que se hizo desde el vocablo griego hasta el inglés. Algo se ha perdido, y su ausencia se hace notar. Un núcleo vacío final no parece suficiente para legitimar la estructura:

(7)

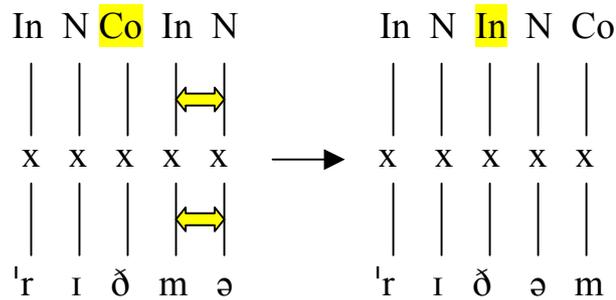


Tomamos esta forma como estructura intermedia, de transición, que es la que busca legitimarse y mueve a encontrar una solución. La diferencia crucial con una inserción a secas es que ahora el N vacío proporciona un espacio en el que una vocal epentética puede **anidar** ⁶.

Estamos planteando la inserción de una schwa porque no es plausible que la schwa estuviera ya allí. O si estaba, también era una vocal inestable en esa posición. Al fin y al cabo, las schwas en inglés provienen de la reducción de una vocal plena. En *dogma* o *asthma* se puede rastrear cuál es la vocal plena de origen, pero ¿qué hay de ella en *chasm* o *rhythm*? Un inconveniente sobrevenido es que con la schwa al final se formaría una palabra muy poco inglesa, /'rɪðmə/. La configuración de esta palabra estaría muy bien para el francés y podría tomarse prestada sin mayor problema, pero habría que importarla en bloque, {rɪðmə}. Lo que no se admite es formar una palabra así desde dentro de la lengua inglesa, con el aparato fonológico propio. Este es un nuevo ejemplo de mecanismos cognitivos que intervienen en la operación de reglas fonológicas.

Lo que se puede hacer entonces es trasladar la schwa a un lugar más “cómodo”. Por medio de una **metátesis** invertimos la posición relativa de /m/ y schwa (y de paso hay que resilabear el rótulo de nodo de la coda, que pasa a convertirse en inicio):

(8)



Una metátesis de este tipo debe de ser la responsable de las CS propias de los auxiliares contraídos, del tipo *haven't* o *didn't*, que tenían la schwa en el núcleo siguiente a la CS, y ahora lo tienen en el núcleo previo, como es habitual en las CS inglesas. Por lo tanto, la metátesis de schwa es un mecanismo atestiguado en otras palabras del inglés, y además, en relación con la formación de CS.

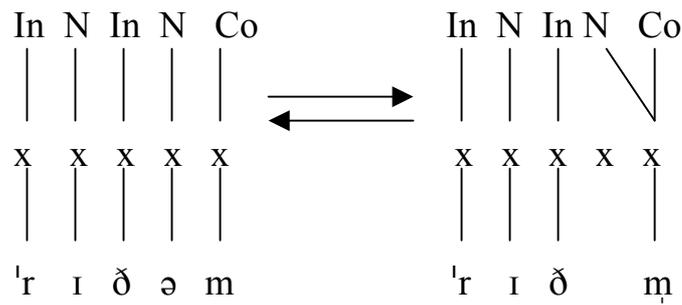
El proceso es un tanto laborioso, cómo no, pero precisamente por ello es tan caro y sólo se recurre a él después de haber agotado otras vías. Además, la forma que se obtiene es un tanto anómala desde el punto de vista del reconocimiento, porque las consonantes que antes formaban un grupo, [ðm] en *rhythm* o [sm] en *chasm*, ahora quedan separadas por la schwa que se interpone. En esta situación, la /-m/ se queda un tanto varada u orillada, y parece fuera de lugar.

Vamos a considerar un ejemplo que parece reforzar esta hipótesis. En la palabra *asthma* se optó por la solución del latín (y el español): eliminar la consonante intermedia y conservar la vocal final. El resultado es /'æsmə/. Pero si hacemos un poco de fonología-ficción, podemos imaginar una evolución distinta, donde la forma griega

asthma pasara al inglés conservando la consonante <th> y eliminando la vocal final (esto es, /'æsðm/ o /'æsθm/). ¿Qué hubiera sucedido entonces? Pues que el resultado habría sido excesivamente complicado. Si insertáramos una schwa, la nueva forma /'æsðəm/ sería francamente irreconocible, porque ahora transmite la impresión de dos sílabas totalmente independientes, (æs + ðəm). Por esto es muy preferible la opción de /'æsmə/.

Nos queda claro que la inserción de la schwa dificulta un poco el reconocimiento de la forma de base o subyacente, puesto que separa las consonantes que en origen estaban juntas. Es aquí donde entra en juego la CS, que no parecía tener papel alguno en la legitimación de una estructura que ya está legitimada. Lo que hace la CS es permitir que las dos consonantes que estaban juntas en la base se mantengan juntas, y con ello facilita el reconocimiento de la forma original. Es decir, que el sistema resuelve la tensión (cognitiva y estructural) a través de un mecanismo que el propio sistema ya tenía, y que no es otro que la formación de CS ⁷. El *cómo* es igual que en las CS que ya conocemos: una consonante se asocia al N previo, sin perder la asociación con su nodo de origen. Llegamos así a la oscilación de estructuras propia de las CS débilmente silábicas, con la clásica alternancia entre /əR/ y /R/:

(9)



Esta oscilación entre las dos formas parece la más adecuada para transmitir, en la medida de lo posible, la estructura básica /rɪðm(ə)/.⁸

Pero la estructura no lo es todo, seguramente. Presumimos que en conjunción con ella operan una serie de factores cognitivos, que sirven de coadyuvantes en la formación de las CS atípicas. Ya hemos mencionado la importancia de la **analogía** con las CS ya existentes en la lengua como motor para ampliar el repertorio a otro campo. Y vamos a añadir ahora una serie de factores que agrupamos bajo el título de **'factores subjetivos'**, dado que dependen mucho de la percepción que se tenga de las cosas en cada comunidad de hablantes y en cada momento histórico de la lengua; como tales, están muy sujetos a variación⁹.

Entre los factores subjetivos incluimos la **hipercorrección**, que hace que los hablantes busquen una pronunciación lo más parecida posible a la ortografía¹⁰. Esta no es sino una variante del reconocimiento de formas, que esta vez toma la forma escrita como base. Otro factor más, que podemos llamar **nacional** (a falta de mejor término), es la percepción que se tiene de las CS dentro de la lengua inglesa. En concreto, nos da la impresión de que las CS se perciben como “muy

nativas” por parte de los hablantes, de modo que las CS se consideran como algo que da sabor inglés a la lengua, y por lo tanto se favorece su uso a pesar de que resulten un tanto “caras” de formar. Este factor se combina con la **percepción estética** que se tenga de las CS, que acaso se aprecien como bellas o atractivas por parte del hablante nativo.

La influencia de estos factores subjetivos sobre la economía fonológica no debe minusvalorarse, en nuestra humilde opinión. Su peso puede justificarse por qué se opta por una solución estructuralmente “cara”, dada la compensación cognitiva que aporta. Esto es, las reglas fonológicas encuentran **motivación** y también **restricción** por parte del componente cognitivo del cerebro.

Hemos planteado cómo las CS atípicas responden a la necesidad del componente cognitivo de facilitar el reconocimiento de formas, y cómo se recurre a la analogía con las CS ya existentes. Pero aún nos atrevemos a dar un paso más para sugerir que las CS sean la **forma primaria** de la oscilación. Nuestra conjetura es que la analogía (mezclada con hipercorrección) dé lugar directamente a la CS, dejando la forma con schwa como derivada. Gráficamente, el proceso es:

(10)

$$\mathfrak{R} \leftarrow \mathfrak{aR}$$

Se trata ni más ni menos que de un proceso de **regresión** o **derivación regresiva** (*back formation*),¹¹ donde no hay una inserción de schwa al uso, sino que se origina a posteriori, mediante una regla que se da por

aplicada. La CS se origina en el plano silábico, mediante la regla de inventario que ya tiene la gramática del inglés; y al ascender al plano métrico, la presencia de la CS desencadena la **reconstrucción** de una forma con schwa, volviendo sobre los pasos de una “presunta” regla de reducción. Naturalmente, la forma con schwa no permite llegar hasta la base, hasta una vocal plena que nunca ha existido, y queda un poco “en el aire”. Pero sí que se da lugar a una alternancia propia de las CS débilmente silábicas en la zona más alta de la estructura.

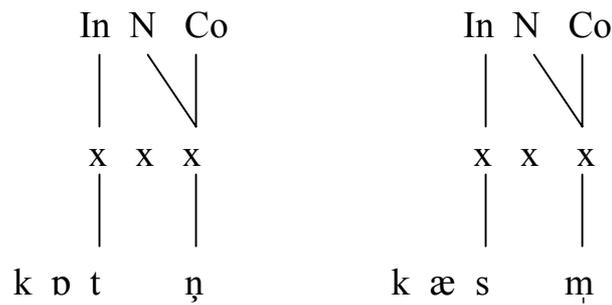
Dejamos planteada así una segunda hipótesis con respecto a la formación de CS atípicas, y lo cierto es que nos gusta más que la epéntesis, que resultaba tan cara y con riesgo de sobregenerar. A favor de la epéntesis juega la analogía con los auxiliares contraídos (*haven't, didn't*), que tuvieron que formarse por metátesis, por lo que es un mecanismo atestiguado en la lengua. Pero si nos gusta más la hipótesis de la ‘regresión’ es porque tiene la ventaja de no postular el manejo de schwas en la zona baja de la estructura. Si acaso, habría que recurrir a una schwa léxica, archivada por analogía con las schwas métricas, para optar a la inserción (una vez más, sería el mecanismo de la analogía en juego). Pero nosotros nos quedamos con la idea de que la schwa sólo se forma por semitransparencia en el filtrado de rasgos al plano métrico, y preferimos evitarla por debajo de este nivel. En inglés, la schwa debería ser propia, casi exclusiva, de la parte débil de un pie, siempre dentro del ámbito estricto del plano superior¹².

CS típicas y CS atípicas

Terminamos este capítulo comparando las CS típicas y las atípicas, gracias a lo cual profundizamos en la comprensión de los dos tipos de CS. Hemos señalado que el *porqué* de su formación es distinto: que las CS típicas se forman dentro de la estrategia de primer y segundo plano que favorece el procesamiento de la información. Y que en las CS atípicas hay un problema de origen fonológico (la imposibilidad de silabear una secuencia melódica, por incompatibilidad con el Principio de Secuenciación Sonora), que mueve a buscar soluciones. Cuando está disponible una solución más “barata” o económica, esa es la que se adopta (la coalescencia de consonantes en *phlegm*). Y de no ser así, hay que recurrir a una solución más “cara”, que es la inserción de una vocal fría sobre un núcleo vacío, con metátesis y con alternancia con CS. O bien se forma una CS para legitimar la secuencia, y por regresión se reconstruye la variante con schwa. Así es como un mecanismo que el sistema ya tenía (para facilitar el procesamiento métrico, cognitivo) se **reutiliza** en otro aspecto (también desde el componente cognitivo, pero con otro fin, que es facilitar el reconocimiento de la forma original, subyacente).

Pese a la diferencia en el *porqué*, el *cómo* es esencialmente igual en los dos tipos de CS. El mecanismo para formar la CS típicas se traslada sin más a las atípicas, de modo que se mantiene la asociación de una consonante al núcleo precedente:

(11)



Es importante sin embargo la diferencia de ubicación que separa a estos dos tipos de CS, y que viene de la mano de su diferente motivación. Con un mismo *cómo*, un distinto *porqué* acarrea un distinto *dónde*. Las CS típicas se forman, de acuerdo con su motivación métrica, en el plano más alto. Es necesario que esté formado el pie trocaico para que opere el mecanismo de su formación.

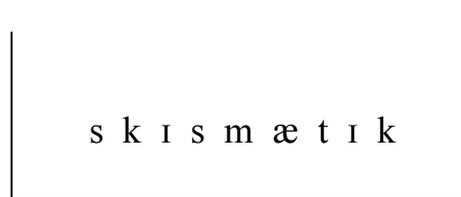
Por su parte, las CS atípicas tienen que formarse en la transición desde el componente melódico al plano silábico. Una consecuencia relevante es que en este ámbito no hay captura de mora (que es cosa del plano superior). Por el contrario, es regla general que en el plano silábico se maximicen los inicios, y esto es lo que vamos a encontrar en las CS atípicas. Buena prueba de esto es que en las formas con schwa de *citoplasm* o *minimalism* no vemos un cambio de la /s/ hacia su equivalente sonora, la /z/. Así suele ocurrir cuando la /s/ está entre vocales (el llamado *s-Voicing*, donde /s/ → /z/ /V__V),¹³ y así lo vemos en *bosom* /'bʊzəm/, que sí que tiene /z/. Esta sonorización se hace en el ámbito de un troqueo, con toda probabilidad, asociado a una captura de mora (véase Harris 2004). En cambio, la /s/ insonora de *chasm* /'kæsəm/ nos hace pensar que está silabeada como [kæ]·[səm], donde la /s/ ocupa el inicio de la sílaba recién formada, y

así se va a quedar para evitar problemas de reconocimiento, como ya hemos dicho.

Hay un último detalle de suma importancia que no queremos pasar por alto. Las formas derivadas con vocal, del tipo *cytoplasmatic* o *rhythmic* nunca admiten ni la CS ni tampoco la schwa. Jamás. Si ya es imposible oír *schismatic* como */skɪs¹mætɪk/ (donde la CS portaría el acento de palabra, extrañamente), aún más imposible resulta */skɪsə¹mætɪk/, con la schwa. Una vez más, la clave está en el *dónde*, que ahora es también el *cuándo*.

Asumimos que el afixo viene ya incluido en la secuencia del plano melódico:

(12)



y como tal se silabea sin problemas al proyectar al plano silábico. En cambio, la secuencia s-k-ɪ-s-m necesita de la schwa y la CS al proyectar al plano silábico, y es en esa transición donde se forma. Para entonces, *schismatic* ya está silabeado, y la schwa o la CS ya no encuentran lugar.

Se puede traducir el problema a un sistema de **reglas ordenadas** clásico, donde la regla de sufijación se aplica antes de la de formar la CS o de insertar la schwa. Y diríamos que la sufijación **sangra** la aplicación de las otras dos reglas:

(13)



A sufijación

B insertar schwa, formar CS

Pero insistimos: se aplica antes porque el ámbito donde se aplica está antes. Al estilo de la Fonología Cíclica, se delimita un ámbito de aplicación de la regla, de manera que una vez que pasa, ya no ha lugar¹⁴. En conclusión, el lugar donde se forman las estructuras o formas determina o condiciona el orden de reglas. Una vez más, el *cómo* implica el *dónde*.

Concluimos este capítulo con la satisfacción y la sorpresa de haber encontrado un nuevo porqué al final de nuestro estudio. Un nuevo tipo de CS que se basa en la ampliación de los recursos ya existentes. Y con él hemos encontrado además un nuevo ámbito del mapa fonológico para formar CS. Este es el panorama completo de las CS en la lengua inglesa, y ha resultado más amplio y más rico gracias a la conjunción de formas típicas y atípicas.

Y más allá de las CS y de la propia fonología en sí, nos parece que este modelo tiene consecuencias de largo alcance. La manera en que se aplica un mismo mecanismo en lugares distintos de la arquitectura fonológica guarda un parecido sustancial con la implementación del material genético para producir el diseño de los seres vivos. Los avances recientes en **Genética Evolutiva del Desarrollo** (*evo-devo*) han revelado cómo la evolución de los seres vivos opera en bloques, mediante la recombinación o reaplicación de mecanismos complejos

ya existentes¹⁵. Entendemos que el diseño de la facultad del lenguaje, que al fin y al cabo es una estructura biológica, pudiera seguir estas mismas pautas, y obtener así una gran complejidad con mecanismos sencillos, según la aspiración del Programa Minimista.

¹ Nótese que aquí no hay un núcleo vacío que permita licenciar una ‘estructura africada’ como la que vimos para las CS típicas. Aparte de su interés intrínseco, las CS atípicas tienen el interés añadido de que nos sirven para corroborar hipótesis que hemos formulado previamente para las CS típicas.

² El argumento reafirma su validez con los derivados de *diaphragm* y *apothegm*. En cualquier caso, la evolución de *phlegm* y otros derivados no deja de ser curiosa, y da para mucho estudio. Por ejemplo, en inglés *phlegm* tiene un derivado “sincrónico” *phlegmy* /'flemɪ/, y otro “diacrónico” *phlegmatic* /fleg'mætɪk/.

También hay que contar con la evolución en francés (desde donde se prestó al inglés), porque hubo una coalescencia de la vocal con la /g/: *flegma* → *fleume*, donde /e + g/ → /ø/. No pensamos que la coalescencia de la vocal se haya trasladado al inglés, en vista del derivado *phlegmatic*, con /g/, y también de *diaphragm* o *apothegm*. Pero sí que sirve como argumento a favor del mecanismo de coalescencia (véase *infra*, en el cuerpo del texto).

³ Queda pendiente de explicar qué mecanismo se aplica en *sign*. Asumimos que también aquí hay una /g/ subyacente, en vista de los derivados *signal* y *signature*. Pero en /sain/ se ha optado por eliminar la /g/, y se ha hecho a pesar de no tener la relación de afinidad que acercaba la /g/ y la /m/ en *phlegm*. ¿Cómo así? Tenemos dos hipótesis al respecto, que no son incompatibles. La primera es que por **analogía** se haya trasladado la coalescencia de /g + nasal/ a todas las secuencias de este tipo. Y la segunda es que haya una coalescencia /i + g/ similar a la que da lugar al alargamiento compensatorio en *sigh* o *night*. Tomamos como indicio de esta coalescencia /i + g/ el alargamiento de la vocal /i/ (con diptongo subsiguiente, /i:/ → /ai/), alargamiento que en cambio no se produce en *phlegm*, *diaphragm* o *apothegm*.

⁴ Las soluciones que se adoptaron al incorporar los préstamos resultan ser variopintas, porque en un buen número de casos se optó por conservar la /-a/ final. Por ejemplo, en *dogma*, *stigma*, *magma*, *enigma*, y en *miasma*, *plasma*. Pudiera pensarse que el registro religioso tenga influencia en *dogma* o *stigma*, para que el préstamo sea más conservador, pero *schism* contradice esta hipótesis.

Aún se complica más la cuestión con *theme* o *scheme*, donde parece haberse aplicado una regla de MEOSL (*Middle English Open Syllable Lengthening*) [recuérdese el capítulo 5], que sin embargo no se aplicó en *problem*. Probablemente la estructura tiene que ver: el troqueo que hay en origen para *theme* y *scheme* es un ámbito de aplicación del MEOSL, pero no sucede así para *problem(e)*.

Y en el caso de *phlegm*, parece ser que el alto coste de la formación de una CS fue el motivo para no formarla. Pero era perfectamente posible, y además sería casi una ‘antigeminoide’, dada la afinidad entre las dos consonantes.

Estamos convencidos de que hay también factores cognitivos involucrados, y de que el patrón de incorporación de préstamos dice mucho sobre la economía del sistema, tanto en lo fonológico como en lo cognitivo. Lamentablemente, el tema es demasiado amplio y escapa al campo de

investigación de esta tesis. Pero estamos seguros de que será un terreno fértil para nuevas investigaciones.

⁵ En inglés sólo tenemos en mente la posible inserción en los plurales de palabras terminadas en sibilante, como *buses*, *wishes* o *churches*.

En holandés es más evidente la epéntesis de schwa. Es un proceso opcional y se lleva a cabo entre una líquida y la C siguiente: por ejemplo, *help* ('ayuda') puede ser tanto /help/ como /heləp/, *worp* ('lanzar') es [wɔrp] o [wɔrəp], *balk* ('rayo') es [balk] o [balək], y *snurk* ('roncar') es [snɔrk] o [snɔrək].

Pero el proceso está sujeto a restricciones: la inserción no es posible cuando la obstruyente final es una coronal. Por ejemplo, en *vilt* ('sintió'), *hard* ('corazón'), *hars* ('resina') o *arts* ('médico') no es posible la schwa.

Curiosamente, la inserción de schwa siempre es posible cuando la consonante final es una sonante (incluso si es coronal). Así, *helm* ('casco'), *arm* ('brazo') y *kern* ('pepita', 'grano') admiten la schwa sin problemas.

Y más aún, la schwa es obligatoria cuando la distancia sonora es demasiado corta. Concretamente en -*rl (el equivalente de *Karl* (en alemán, Carlos) y *Charles* (en inglés) es *Karel*. Y lo que en alemán es *Kerl* ('patán', 'palurdo') o en inglés es *churl*, en holandés es *kerel*. Tomado de Ewen & van der Hulst (2000: 190-93).

⁶ Que sepamos, esta idea de anidamiento (*landing*, *landing site*) propia de la sintaxis (de la sintaxis generativista, se entiende: véase Haegeman (1991) o Radford (1988) para más detalles), es nueva en la fonología. Los núcleos vacíos de la FR nunca se habían planteado con esta idea, según nuestro mejor entendimiento.

⁷ Retomando el ejemplo de *asthma*, señalemos que en este caso ni siquiera la CS bastaría para volver a juntar las consonantes separadas, puesto que /'æsðm/ sigue dando la misma impresión de (æ + ðm), que pierde de vista por completo a la forma original, (asð + ma). Por eso tampoco se adoptó esta solución con *asthma*, y se prefirió *asma*.

⁸ La verdad es que se ha tomado con mucha naturalidad la presencia de un mecanismo reparador en un sistema de reglas (esta es la hipótesis que ha seguido la Fonología de la Rección para explicar las CS en general), cuando parece que debería corresponder más a un sistema de restricciones. Al fin y al cabo, ¿puede una regla que ya está incorporada a la gramática producir secuencias mal formadas? En principio, no, pensamos. Pero la necesidad de reparación proviene de las exigencias de otro sistema, de otro ámbito distinto de aquel en el que está operando la regla. Por ejemplo, si producimos una secuencia {a-g-m}, puede ser que incumpla el Principio de Secuenciación Sonora, pero desde el punto de vista melódico es impecable, ya que está formada por segmentos propios del léxico de esa lengua. La secuencia está bien formada para el componente melódico.

Por otro lado, al tomar prestado material de otras lenguas, es fácil que haya que "reparar" o adaptar. Y todas las lenguas están en evolución permanente (por definición, es intrínseco al lenguaje). Todo esto no se podría llevar a cabo sin mecanismos reparadores o adaptativos. Y también hay que contar como "reglas imperfectas" todas las posibles "infidelidades" que se produzcan durante la adquisición del lenguaje, cuando hay que salvar la discontinuidad de transferir una lengua de un hablante a otro.

Con toda esta perspectiva en mente, nos declaramos partidarios de una gramática que tenga tanto reglas como restricciones, o mecanismos de reparación, de adaptación, y también de "control de la calidad" (como se hace en las vías del sistema nervioso central con las estaciones de refresco y el *feedback* desde puntos más periféricos de la eferencia).

⁹ Al tener en cuenta el factor subjetivo, podemos acercar posturas con la **Lingüística Cognitiva**. Por ejemplo, Romano (2001: 399) plantea la importancia de la subjetivización:

(...) todos los procesos de cambio semántico descritos más arriba, incluidas la metáfora y la metonimia, se pueden enmarcar dentro de lo que se ha denominado la hipótesis de la subjetivización o la tendencia general que siguen algunos fenómenos lingüísticos hacia la interiorización, es decir, hacia la expresión de la actitud, la implicación o el punto de vista del hablante o emisor en la forma lingüística de su enunciado (Langacker 1990; Traugott, 1989, 1990, 1995, 1996).

Las CS atípicas muestran cómo la subjetivización es un factor significativo, pero al mismo tiempo necesita de unos factores estructurales que la hagan posible. Por eso hemos articulado este trabajo a la luz de una interacción entre cognición y estructura. Al igual que la evolución de las especies biológicas está condicionada por el medio, pero requiere de la maquinaria celular para salir adelante (véase Sampedro 2002). Con esto manifestamos nuevamente nuestra esperanza de que el enfoque paralelo al de la biología ‘evo-devo’ traiga buenos resultados al campo de la lingüística.

¹⁰ En fecha reciente, un hablante nativo estadounidense, de nivel cultural alto, nos trasladaba su irritación por lo que él consideraba dejadez en la pronunciación del inglés. Concretamente, se quejaba de que no se pronunciara la /g/ en los gerundios ‘-ing’ y otras formas.

Una forma de hipercorrección basada en la ortografía es la que cita Cruttenden (2001: 113): en dialectos del norte de Inglaterra, la <u> de *sugar*, *butcher*, *cushion* se pronuncia /ʌ/.

¹¹ El nombre de ‘derivación regresiva’ es la traducción de *back-formation* que se sugiere en Alcaraz (2000: 53), según Cruz Cabanillas (2001: 711), *ya que se crea una nueva palabra mediante el acortamiento de otra existente para recuperar una supuesta base léxica que nunca se atestiguó en la lengua*.

En inglés hay varios verbos “denominales” formados por este método: *burglar* < *burgle*, *liaison* < *liaise*, *surveillance* < *surveil*, *peddler* < *peddle*, *editor* < *edit*, *incentive* < *incent* (Lieber 2010: 51-52). Y las formas “en singular” *pea* < *pease*, o *cherry* < *cherise* (Algeo: 2010: 238), entre otras, son ejemplos muy conocidos de derivación regresiva en inglés.

Una vez más, el factor subjetivo, cognitivo, interviene a la hora de crear o de atribuir reglas y derivaciones. Creemos que la derivación regresiva es otro ejemplo, junto con la etimología popular de cómo actúa la fonología receptiva (véase el capítulo 10 para este concepto).

¹² Aunque hay que tener en cuenta la posible inserción PCO (Principio del Contorno Obligatorio) de *buses*, *wishes*, *churches*, claro. Por eso no podemos darle la exclusividad absoluta al plano métrico.

¹³ Roca & Johnson (1999: 543ss.) es muy ilustrativo.

¹⁴ Recomendamos Roca & Johnson (1999) para una explicación esclarecedora de estos aspectos teóricos.

¹⁵ Recordemos la idea de ‘evolución modular’ de Sampedro (2002) que recogimos en el capítulo de introducción de esta tesis.



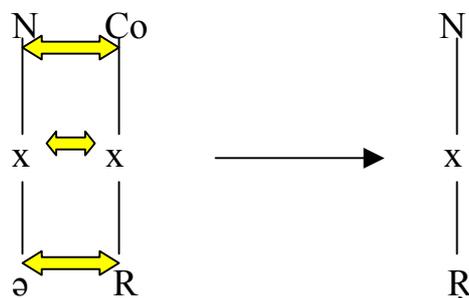
12 Cómo se forman las CS en checo

Estructura de las CS checas

Todo nuestro estudio hasta el momento se ha centrado en las CS del inglés, tratando de no dispersar excesivamente nuestro esfuerzo. Pero si nos limitamos a una sola lengua, corremos el riesgo de tener una perspectiva demasiado estrecha y de que nuestra teoría pierda valor explicativo. Para mejorar nuestras intuiciones sobre las propias CS del inglés, ensancharemos el enfoque y trasladaremos nuestra teoría a las CS del checo (o eslavas). Así veremos en qué difieren de las inglesas (o germánicas, más generalmente), sin dejar de lado los puntos de contacto entre lenguas. Y, en sentido contrario, esperamos que el modelo teórico y todas las explicaciones que hemos desarrollado demuestren su utilidad para abordar el estudio de otras CS.

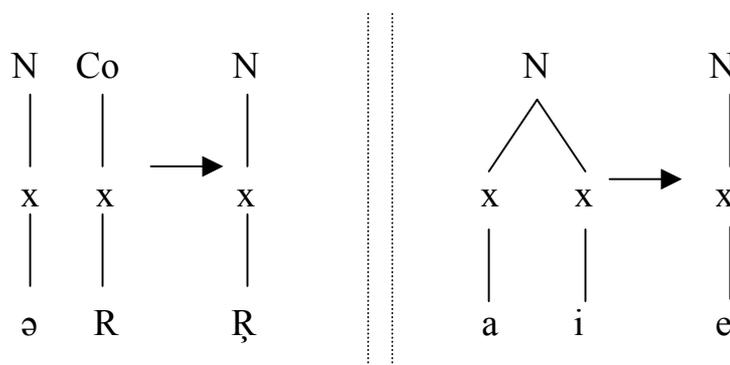
Con el mismo modelo de siempre, proponemos que las CS checas se forman igualmente por coalescencia de una schwa y una sonante en torno a un nodo N. Ahora bien, la estructura que se forma en checo no es idéntica a la de las CS inglesas, sino que tiene algunas características propias. Esta vez postulamos que la coalescencia es total, hasta el punto de abarcar también a los nodos silábicos. La regla de formación de CS en checo debe ser, por tanto:

(1)



Se trata de una variante del esquema que ya conocemos. Como de costumbre, partimos de un N ocupado por una schwa, que por coalescencia se condensa con la consonante de la coda. La condensación de segmentos produce una consonante asociada a un nodo N, y además también afecta a las unidades de tiempo. Pero en las CS checas, según refleja el diagrama, la condensación llega también a los nodos silábicos, de modo que sólo queda un nodo N al final del proceso. Es una coalescencia más profunda, que deja un único nodo silábico (y un solo segmento y una sola unidad de tiempo), como sucede en la coalescencia de diptongos. La similitud de estas dos estructuras se aprecia mejor con un diagrama:

(2)



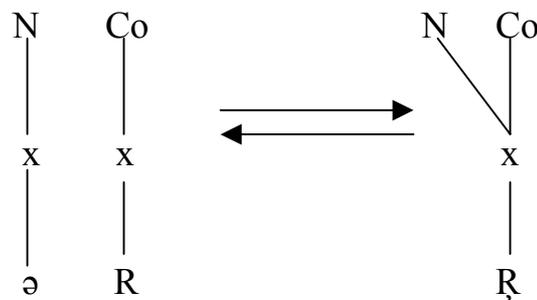
La estructura de las CS checas remeda la que proponía Zoe Toft (2002) para la /ɹ/, pero tiene una diferencia crucial, porque en este

esquema no hay inserción directa, sino que es el paso final de todo un proceso con varias representaciones intermedias. Nos parece más justificado hacerlo así que recurrir a una inserción directa de una consonante en el nodo N, que a nuestro entender resultaría artificial y muy poco restringido.

Otra característica diferencial de las CS checas es que su proceso de formación no es reversible, y por eso lo hemos representado con la flecha de un solo sentido. Nos parece muy probable que las dos características estén relacionadas. Que al suprimirse la coda por la condensación de nodos, la estructura no puede oscilar. Porque la coda no está disponible para que la R retorne a ese nodo, y si queremos revertir la formación, tenemos que revertir la regla, con todo el proceso en su conjunto; lo cual no es imposible pero sí que resulta más trabajoso.

Irreversibilidad y condensación de nodos distinguen a las CS checas de las inglesas, que en los demás aspectos coinciden. Recordemos el esquema de formación de las CS inglesas para establecer mejor la comparación:

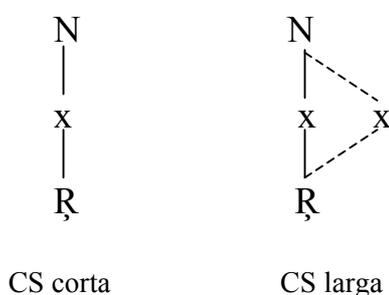
(3)



El carácter reversible del proceso en inglés permite a esas CS alternar con una schwa, como hacen tan a menudo. Son las CS débilmente silábicas típicas del inglés, del tipo de la que vemos en *cotton*, que puede ser tanto /'kɒtŋ/ como /'kɒtən/. Por el contrario, el proceso **irreversible** del checo no deja posibilidad de volver atrás. La regla conduce a un final de camino, que resulta en una CS plenamente formada y que es la única estructura válida. De modo que las CS checas siempre son **fuertemente silábicas**. La regla irreversible manda que no alternen con la forma previa, la que tiene la schwa, y así se cumple en la práctica; el esquema teórico propuesto parece estar haciendo predicciones correctas.

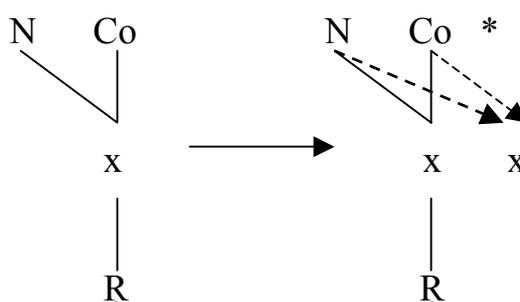
Otra consecuencia del modelo teórico es que las CS eslavas son susceptibles de un **alargamiento subsiguiente**¹. No es una mera posibilidad teórica, porque el eslovaco distingue entre CS largas y cortas. En Blaho (2004: 50) vemos ejemplos concretos: su diagrama (41) compara *srna* ‘rebeco o corzo hembra’, *prst* ‘dedo’, *vlna* ‘ola’, *čln* ‘barco’ con *trň* ‘espina’, *krč* ‘calambre’, *kl'znuť* ‘deslizamiento’, *kl'b* ‘articulación’.²

(4)



planteado. También es interesante el contraste con el inglés. Aquí no se pueden formar CS largas, ya que están asociadas a dos nodos silábicos a la vez, y al extender su longitud violaríamos un principio fundamental de buena formación de la Fonología Autosegmental: la **condición de no-entrecruzarse** (Goldsmith 1976: 146)³:

(5)



En definitiva, la coalescencia total del eslavo produce una diferencia de estructura. La diferencia es determinante. Las CS del checo han perdido el nodo de coda y sólo les queda un nodo N, sin ramificación. Tienen capacidad para la elongación, pero en cambio no pueden oscilar entre nodos, y tampoco podrán ser ambisilábicas, como veremos.

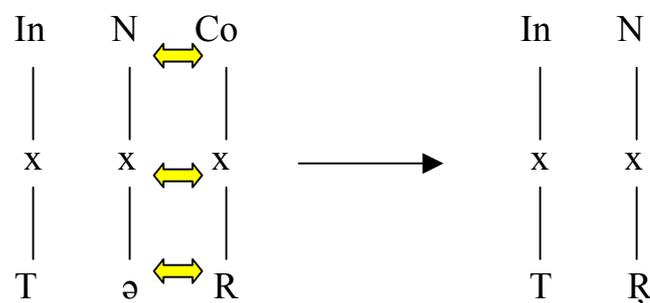
Dónde y cómo se forman las CS checas

Las CS inglesas y las checas difieren de manera esencial en el lugar donde se forman (recordemos que el *cómo* guarda relación estrecha con el *dónde*). Las CS inglesas obedecen a un motivo métrico, y por eso se forman en el plano métrico, en la zona más alta de la estructura. Pero las CS checas no están asociadas a una posición específica de los pies métricos, y de ahí que nos inclinemos a localizar su formación

por debajo del plano métrico. Ya que los nodos silábicos confluyen en uno solo ($N + Co \rightarrow N$), las CS tampoco pueden estar por debajo del plano silábico, y por esto es por lo que ubicamos su formación en este plano de la estructura fonológica.

Casi siempre hemos utilizado la representación al uso, con nodos y segmentos, y en general la vamos a seguir empleando. Nos gusta y creemos que da buenos resultados: por ejemplo, ha sido muy gráfica con las CS largas del eslovaco (y su imposibilidad teórica en inglés). A modo de referencia, vamos a repetir cómo se representa en este sistema gráfico la formación de CS en checo:

(6)



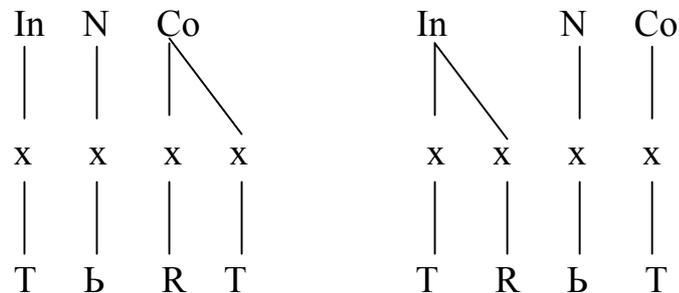
Pero este tipo de esquema también puede dar lugar a confusión, porque se combinan los segmentos del plano melódico con los rótulos de nodo del plano silábico. Para evitar el malentendido, vamos a recurrir una vez más a la **Estructura de Frase Escueta** (*Bare Phrase Structure*), donde los elementos de un plano se proyectan al plano siguiente. Empezamos por el plano melódico, y ahí sólo encontramos una secuencia de segmentos, que están meramente enunciados:

(7)



La presencia de la schwa es fundamental. Sospechamos que siempre tiene que haber una schwa subyacente para que se formen las CS en checo. La schwa en este caso representa una vocal central, que de acuerdo con Scheer (2009) se deriva de la *yer* eslava, una vocal histórica que se representa como <ɐ>. Scheer (2009) afirma que las CS checas provienen de dos tipos de secuencias: tanto TBRT como TRBT. Como es natural, las estructuras silábicas son diferentes para cada una:

(8)



A pesar de las dos estructuras y secuencias distintas, nosotros nos decantamos por una sola regla para formar CS. Sea en checo o en otras lenguas, las CS siempre se forman por:

(9)



En inglés es fácil demostrar que la regla es siempre así, dado que la alternancia con schwa muestra el proceso bien a las claras. Para el eslavo no tenemos una forma abierta (con schwa), pero sí que hemos encontrado un indicio favorable a nuestra hipótesis. Está en el himno serbio, en el estribillo:

(10)

Bože spasi, Bože hrani
Srpske zemlje, srpski rod.

Al cantar la melodía, *srpske* tiene una melodía silábica, y se escucha una Ṛ normal y corriente. Pero la CS de *srpski* se canta en dos notas, lo cual obliga a alargar la presunta CS, y entonces se oye con toda nitidez una vocal. Así podemos identificar su naturaleza: es una vocal ‘fría’, central, que podemos etiquetar como schwa y que postulamos como vocal subyacente de la CS. Y además se pone de manifiesto que esa vocal está antes de la CS. La CS en dos notas es, a nuestro parecer, una versión exagerada del patrón de todas las CS eslavas (y de cualquier lengua). Generalizamos el razonamiento y concluimos que las CS, tanto en inglés como en eslavo, son siempre (ə + R). Las CS no son nunca (R + ə), porque ese orden corresponde a las consonantes atrapadas, y en checo o en inglés no hay tal cosa, sino sólo CS, que por definición son post-nucleares⁴.

En conclusión, se llega a un destino común desde orígenes distintos. Hay que confluir en un orden (ə + R). Por consiguiente, las secuencias TṚḄT tuvieron que sufrir una **metátesis** o **reanálisis** para poder confluir con las secuencias ṬḄRT, que son la única secuencia apta para formar CS. En favor de una posible metátesis en checo, podemos

aducir que también se produce en inglés: al menos con los auxiliares contraídos *didn't* y *haven't*, y acaso también con las CS atípicas de *chasm* y *rhythm* (donde el N que se queda vacío al principio está a la derecha de la CS, y después la schwa debe colocarse a la izquierda de la CS, como veíamos).

Nos disponemos entonces a formar la CS, que toma como base la secuencia /TɔRT/ o /TəRT/. Esta secuencia se enuncia en el plano melódico, y no hay nada en ella que rompa el **Principio de Secuenciación Sonora (PSS)**. Es una diferencia sustancial con lo que encontramos en las CS atípicas del inglés (*chasm*, *rhythm*). Por eso nos resistimos a explicar las CS checas como una reparación de secuencias anómalas, y nos orientamos en busca de otra razón que justifique su presencia en las lenguas eslavas.

Nuestro planteamiento es que se produce una **coalescencia** lisa y llana entre una vocal central, muy “oscura” o “fría”, con la sonante que las sigue. Se funde una vocal muy poco sonora con una consonante altamente sonora (ya que en checo sólo se admiten CS líquidas)⁵. Y ya está. La fusión se basa en la pura afinidad fónica, sin que medie ninguna razón de necesidad. Igual que en la coalescencia de diptongos.

Para sostener esta postura nos apoyamos en una posible distribución complementaria entre vocales oscuras y CS en las lenguas eslavas. Por una parte, hay lenguas que no tienen CS y sí tienen vocal oscura: la vocal /i/, en ruso y en polaco, y la /ɤ/ en búlgaro⁶. Y por otra parte, hay lenguas que sí que tienen CS y no tienen estas vocales oscuras centrales: el checo, el eslovaco y el serbocroata (o BCS).⁷ Así que nos

parece que las CS bien pudieran ser una versión particular de las vocales oscuras del eslavo.

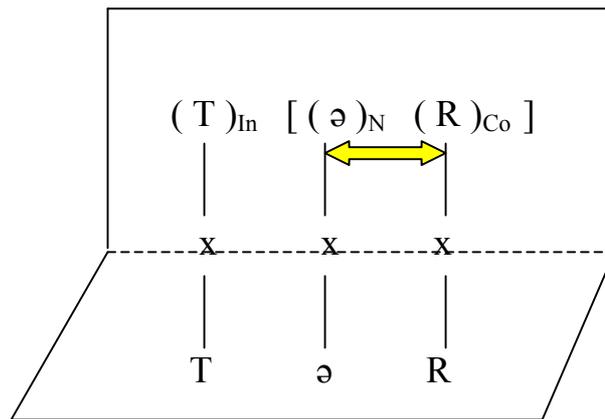
La vocal central forma parte del léxico de la lengua en cuestión. Esto es, asumimos que el checo tiene una schwa o similar en el repertorio léxico (en contraste con la schwa del inglés, que se forma por reducción de una vocal plena en lo más alto de la estructura). La schwa eslava queda en la estructura subyacente (a la vista de lo que se oye en el himno serbio), y así se plasma en el plano melódico. Asumimos también que hay una regla léxica que hace que las vocales oscuras centrales sufran coalescencia automática al entrar en contacto con una líquida.

Dado que la regla de formación de CS en checo sólo cuenta con las schwas que ya estén consignadas en el diccionario, el proceso es poco o nada **productivo**, está **congelado** léxicamente. Muy distinto del inglés, que siempre puede formar nuevas schwas por reducción de vocales en el plano métrico, lo cual permite formar CS en palabras de nuevo cuño y en frases. Es decir, que la congelación léxica, la irreversibilidad del proceso y la falta de oscilación entre nodos (N y Co) hacen de las CS checas un producto muy poco **flexible**, en justo contraste con las CS inglesas, que por todos los motivos opuestos resultan ser bastante flexibles.

La regla que produce la coalescencia tiene que operar en el plano silábico, una vez que los segmentos tienen asignada una posición en la sílaba, porque es ahí donde se establece una relación específica entre los segmentos. La regla de formación de CS exige que la relación entre la schwa y la líquida sea la de un núcleo y su coda. En este

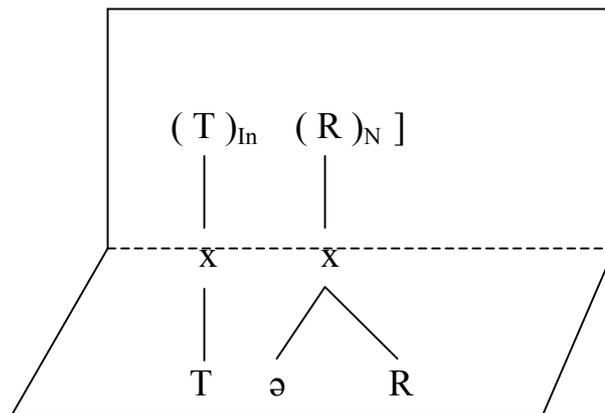
contexto es donde hemos optado por usar la ‘Estructura de Frase Escueta’ (ver arriba), por su capacidad para restringir la operación al plano silábico, y contando con que refleja bien la fusión de los segmentos. Con esta convención representamos primero la proyección desde el plano melódico al silábico, para asignar nodos:

(11)



Y a continuación la formación de la CS por coalescencia, estrictamente dentro del plano silábico:

(12)



donde se conserva la schwa en el plano melódico. Así se recoge en la representación el material de origen y asimismo el producto derivado.

Proyección hacia el plano métrico

Ya tenemos las CS formadas, en el plano silábico, y sólo queda trasladarlas al plano siguiente, que es el métrico. Todo podría reducirse a un mero trámite de proyección, pero las cosas no son tan simples. Las CS checas⁸ tienen una propiedad importante que las diferencia de sus hermanas inglesas: su capacidad (que no la obligación) para portar el acento principal de palabra.

Una circunstancia fascinante, porque para entender la capacidad tónica de las CS checas no basta con mirar al plano métrico directamente, sino que hay que integrar todos los niveles de la estructura. Creemos que el acento se produce a partir de la información que aporta cada plano, y que todos interactúan entre sí. La suma de todos los datos segmentales, silábicos y métricos es lo que hace posible asignar el acento de palabra, a nuestro juicio. Vamos a desarrollar esta postura en los párrafos que siguen.

En el nivel segmental, hay que considerar tanto la vocal como la consonante. Ya vimos en la exposición inicial sobre las CS que el repertorio de CS que tiene cada lengua es paramétrico, de acuerdo con un gradiente de la escala de sonoridad. A diferencia del inglés, el checo limita su repertorio de CS a las líquidas, /l/ y /r/.⁹ Ya que estas consonantes son las más resonantes del espectro, el diferencial de sonoridad con las vocales será el más bajo. Este factor no sólo favorece la coalescencia, sino que además las hace más aptas para portar el acento de palabra que otras consonantes menos resonantes.

Si profundizamos dentro de las clases generales, vemos que la rótica en inglés es una aproximante, /ɹ/, mientras que la del checo es una vibrante /r/. Es posible que la mayor **resonancia** de la vibrante le dé más aptitud para figurar en posiciones tónicas, aptitud que queda fuera del alcance de la aproximante inglesa.

Y en lo que respecta a la vocal, la schwa del inglés es propia y exclusiva de posiciones átonas, por lo que el acento les es inaccesible. Por el contrario, las **yers** (ɐ) sí que podían estar en posición tónica, y por eso no es de extrañar que las CS del checo suelen estar en sílabas tónicas (aunque no exclusivamente, recordemos).

En el nivel silábico, hay diferencias en el grado de identificación de la CS con el nodo N. La fusión de nodos permite a la CS del checo su inserción directa en el nodo N, en tanto que la CS del inglés sólo está asociada al N.¹⁰ En inglés subsiste la unión con el nodo de origen (sea la coda o el inicio), y esto la hace “menos nuclear”, porque su asignación o pertenencia se reparte entre N y Co. Sin embargo, la checa es exclusivamente nuclear. Y como dicen Szigetvári & Scheer (2005), el acento de palabra es una propiedad nuclear, limitado de manera característica a las vocales, que depende de la proyección del nodo N más que del propio segmento en sí.¹¹ Por eso las CS checas son más aptas para portarlo, aunque no debería estar vedado a las CS inglesas, que al fin y al cabo son también nucleares¹².

La existencia de CS tónicas en checo tiene una consecuencia inmediata para la teoría general. Podemos colegir que el acento principal de palabra no busca tanto recaer sobre un segmento (la vocal) cuanto sobre un nodo silábico (el N), sobre la cabeza de este.

Parece que importa más la proyección funcional que el segmento léxico, por así decirlo, a la hora de asignar el acento de palabra. En el modelo ascendente, pues, las proyecciones funcionales son importantes: se proyecta de segmentos a nodos silábicos, de estos a moras y a pies, y de pies a acento. Siempre se tiene la noción de **nuclearidad** o **endocentrismo** (*headedness*) en mente, para una organización de constituyentes como la de cualquier estructura sintáctica bien formada.

En el nivel métrico, las CS inglesas están en la parte débil del troqueo, que por definición no puede recibir el acento principal de palabra porque violaría la **Restricción de Columnas Continuas**.¹³ Y además es muy lógico que la parte débil del troqueo no porte el acento principal de palabra, simplemente desde un punto de vista conceptual. Por su parte, las CS del checo no tienen una asignación métrica determinada, y por eso las encontramos tanto en posición tónica (*vlk*, *krk*, *strce*) como átona (*mohl*, *bobr*)¹⁴.

Resumiendo: las CS inglesas son siempre átonas, mientras que las del checo son aptas para portar el acento principal de palabra, si bien son libres de no hacerlo cuando su asignación métrica así lo dicta.

Las CS checas en su contexto

En el contexto izquierdo, las CS checas se comportan como (casi) cualquier tipo de CS conozcamos por su absoluta necesidad de tener un inicio a su izquierda¹⁵. Lo que es específico de las CS checas es que tienen un mayor contraste de sonoridad con su inicio. El contraste

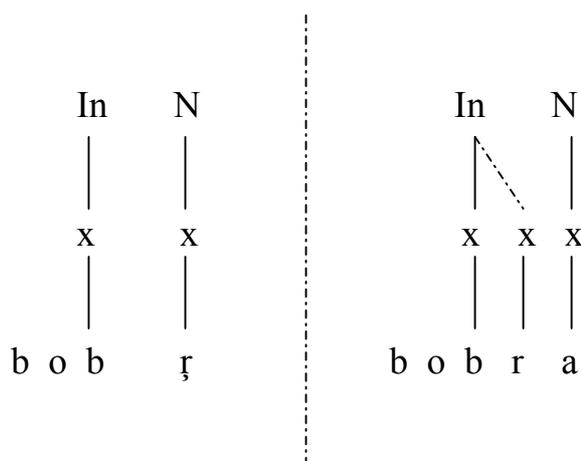
se acentúa tanto por el grado de sonoridad tan alto de la CS (son todas líquidas, no se admiten nasales) como por la tendencia del inicio a ser poco resonante (en checo el inicio de las CS es siempre una obstruyente T, en buena sintonía con la hipótesis TVR). Y al mismo tiempo, esta gran diferencia de sonoridad inicio/CS concurre, curiosamente, con una relación más estrecha entre el inicio y la CS. Debemos tener en cuenta que en checo no existe la captura de mora, frente a lo que sucede en inglés. La ausencia de captura de mora y de CS nasales excluye los cúmulos del tipo /ʔŋ/, tan característicos de las CS inglesas. Cúmulos que por cierto no permiten el menor grado de schwa, y en este sentido son más fuertemente silábicos (a pesar de que alternen con una forma TəR) que cualquier ejemplo del checo, donde todas las CS son de por sí fuertemente silábicas.

En lo que respecta al contexto derecho, es muy definitoria de las CS checas su incapacidad para admitir vocales por la derecha, o sea, *R/ __V. Las CS aceptan el contexto absoluto / __#, o una consonante (generalmente, una obstruyente / __T), tanto en interior de palabra como en posición final. El contacto inmediato puede ser una fricativa, como sucede en *prst*, pero la posición del extremo derecho parece reservarse a sobre todo a las oclusivas¹⁶.

La ausencia de CS seguidas de vocal vuelve a separar al checo del inglés. Algo tiene que ver la diferencia de estructura, y más concretamente, el carácter exclusivamente nuclear de las CS checas, que les impide funcionar como inicio de la sílaba siguiente. Una ventaja que sí disfrutaban las CS del inglés, por su estructura ambisilábica.

También influye el lugar de formación, que para el checo está en el plano silábico (igual que las CS atípicas). En la alternancia flexiva *bobr / bobra* ‘castor’ (NOM sg/GEN sg), la vocal final de *bobra* se ha añadido a la secuencia; cuando el silabeo la asigna automáticamente al núcleo, priva de ese lugar a la ‘R’, que pierde cualquier opción de silabicidad y debe pasar al inicio previo:

(13)



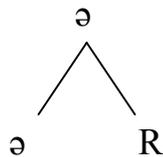
Las CS checas siguen el silabeo canónico, por lo que maximizan el inicio: por eso en *bobra* la líquida ocupa la rama derecha del inicio. Es importante insistir en que la razón de esto es precisamente que las CS checas se forman en el plano silábico, y no en el plano métrico. De nuevo estamos ante un aspecto relacionado con el *dónde* y con el conjunto de la estructura¹⁷.

A la vista de los hechos, asumimos que los sufijos se agregan en el plano melódico, y luego ascienden al plano silábico para silabearse sin más¹⁸. Da la impresión de que las lindes de morfema son transparentes, que no afectan a la formación de CS, como tampoco parecían hacerlo en *bottled* o *evening*, a pesar de que estas CS se forman en el plano métrico.

En la alternancia flexiva entre *bobr* y *bobra* pensamos que siempre está la schwa subyacente en el plano melódico. Esta vocal no puede ascender nunca al plano silábico, porque su proyección es semitransparente, como dijimos. En *bobr* se licencia por la coalescencia con la sonante, y en *bobra* no se forma CS porque la vocal final desbanca a la R del núcleo, y así bloquea este paso.

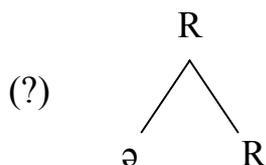
Finalmente, la representación mediante Estructura de Frase Escueta ilustra muy bien la naturaleza de la formación de CS. En condiciones normales, el núcleo está ocupado por una vocal, que es la que proyecta al nivel siguiente (esto es lo que sucede en *bobra*). Cuando la vocal toma un complemento, la estructura es $\{\emptyset, \{\emptyset R\}\}$. Si prescindimos de planos y de unidades de tiempo, para mayor sencillez de la representación, la estructura de la CS es:

(14)



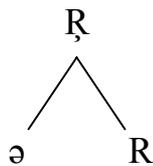
En las CS sabemos que la que proyecta es la ‘R’, pero nos resistimos a que se acepte una anomalía estructural como $\{R, \{\emptyset R\}\}$:

(15)



¿Cómo lo resolvemos? El quid está en que el segmento que proyecta no es exactamente ‘R’, sino ‘R̥’, con los matices que introdujimos en el análisis infrasegmental. A raíz de la coalescencia de rasgos infrasegmentales, se ha formado una ‘R’ silábica, por lo que la proyección no es {R, {ə R}}, sino {R̥, {ə R}}:

(16)



Una diferencia crucial. En lo alto sigue habiendo un nodo vocálico, con los rasgos más fundamentales, que proviene de la schwa (por eso la proyección no es opaca, sino semitransparente). Y luego todos los rasgos que rellenan y dan contenido sonoro a la CS son los rasgos secundarios, que vienen de la ‘R’. Digamos que en esto consiste precisamente una CS: un nodo vocálico, nuclear, pero vacío de contenido sonoro específico, y una matriz sonora que aporta la ‘R’:

(17)

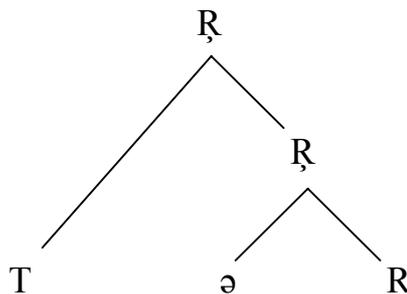


Es la misma idea de **conflación** (*conflation*) que propusieron Hale & Keyser (2000) para la **sintaxis léxica**. El esquema es análogo al de la **incorporación**, pero el elemento incorporado es el único que resulta visible, mientras que el elemento de cabeza queda ‘vacío’, esto es, sin

expresión fonológica. Tenemos, pues, un nodo que llena de contenido la proyección y le aporta su rótulo, y otro asociado que le aporta la matriz sonora.¹⁹

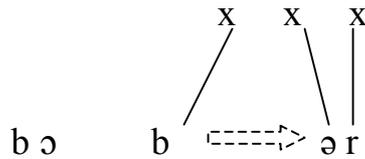
El esquema se puede trasladar sin problemas al inglés, simplemente teniendo en cuenta que en esta lengua las CS se forman en el plano métrico. Pero la coalescencia y la confluencia, en relación con una proyección semitransparente de la schwa siguen siendo válidos igualmente. También conservamos la importancia del inicio (con mayor fuerza si cabe en la lengua checa) para formar una **estructura “africada”** que licencie el conjunto:

(18)



Recordemos asimismo que las afinidades sonoras respectivas ayudan a que la estructura funcione. Por una parte, la schwa y la sonante están muy cerca en la estructura por su gran afinidad sonora (con una mínima distancia sonora, y también con la afinidad del tercer formante, F3, de acuerdo con Ladefoged (2001a: 188)).²⁰ Y al mismo tiempo, el inicio y el núcleo de la CS quedan a gran distancia, dado que en checo el inicio tiene que ser una T (no puede ser una R). Esta diferencia de distancias (correspondientes a las afinidades sonoras), por ejemplo en *bohr*, es:

(19)



Con los ángulos de enlace apreciamos cómo la gran distancia sonora que la obstruyente ('T') del inicio marca con la schwa ayuda a la coalescencia. El inicio literalmente “empuja” a la schwa contra la sonante (indicado con la flecha). Por eso en el diagrama de la estructura “africada” pusimos a la T en la línea de base, para indicar la importancia de las relaciones entre segmentos. La adyacencia inmediata entre T y schwa es la que marca una distancia sonora suficiente para “empujar” a la vocal contra la R. Si por el contrario la T estuviera seguida por la R, la distancia sonora y la diferencia de rasgos ya no sería tan fuerte, y ya no habría coalescencia. La proyección sería entonces $*\{R, \{R \ ə\}\}$, que es claramente agramatical²¹.

Hemos llegado así al final de nuestro estudio de las consonantes silábicas. La aproximación a las CS checas que acabamos de realizar ha puesto a prueba la validez del modelo teórico de esta tesis, a nuestro juicio con buen resultado. Y desde luego que ha proporcionado una visión más amplia y fructífera sobre cómo y por qué se pueden formar CS en una lengua.

¹ Recordemos que hasta ahora el modelo teórico había descartado las CS largas. En eslovaco, como veremos inmediatamente, la elongación es posterior al tercer paso, por lo que no contradice nada de lo dicho hasta ahora.

² En eslovaco, la tilde sobre una vocal (y en este caso, también sobre una CS) indica que un segmento es largo.

³ La restricción que prohíbe el entrecruzamiento de ramas (*no-crossing constraint*) es común a la sintaxis (ver Radford 1988: 120-22), y debe de funcionar como principio universal de la estructura en el lenguaje humano.

⁴ Esta es la definición ‘clásica’ de Scheer. En Scheer (2009), recordemos, las consonantes atrapadas no están asociadas a un N, y las CS ramifican hacia la derecha. En cualquier caso, no se pueden aceptar dos sentidos de ramificación. Si postulamos que las CS son (ə + R), entonces no podemos admitir (R + ə).

⁵ Todo es susceptible de complicarse: hemos visto una /ŋ/ silábica en un himno de la Iglesia Ortodoxa, que está escrito en antiguo eslavo eclesiástico. En la versión cantada por Divna Ljubojevic se oye perfectamente la CS, muy al estilo de las CS del inglés. Nos referimos al verso de la Gran Doxología (*Veliko Slavoslovlje*):

*Slava Tebe pokazavšemu nam svjet
Slava v višnjih Bogu (...)
Gospode Bože, Agnče Božij, Sine Oteč*

⁶ La vocal /i/ se escribe <И/и> en ruso, y en polaco es <y>. La /ɤ/ del búlgaro se escribe <Ь>.

⁷ BCS incluye el bosnio.

En un lugar intermedio puede quedar el esloveno, que en su escritura parece tener CS, pero al parecer al pronunciarlo sí que hay una schwa, que es parte del inventario de vocales de esta lengua. Y cómo no, la schwa está antes de la R.

⁸ Y las del tashliyt, que se han quedado fuera de este estudio.

⁹ La ausencia de CS nasales en checo impide que haya **antigeminoides** como las del germánico (donde la CS nasal asimila el lugar de articulación del inicio, recordemos). Por ejemplo, *bleiben* realizado como [bm].

¹⁰ La fusión en torno a un solo nodo, que además resulta ser el nodo N, probablemente produzca un grado más fuerte de identificación entre la schwa y la sonante, de modo que aumente la sonoridad y favorezca su presencia en la posición tónica. A este efecto, nos gusta la representación que se hace en la fonología de elementos, donde los segmentos melódicos /ə/ y ‘R’ se agrupan en una sola rama.

¹¹ De acuerdo con este punto de vista, el peso que reciben las consonantes de “coda” se debe a que van seguidas por un N vacío.

¹² Dando la vuelta al argumento, recordamos cómo Scheer (2003b) dice que la presencia en posición tónica es una prueba de que las CS están asociadas a un núcleo silábico, a diferencia de las ‘consonantes atrapadas’. Las otras dos diferencias son que las CS cuentan para el verso y que no son transparentes a la sonorización.

¹³ La Restricción de Columnas Continuas (debida a Prince 1983, y recogida en Hayes 1995: 34-37) impide las discontinuidades en las gradillas métricas. La restricción dice que una gradilla que contenga una columna con una marca en la línea $n + 1$ pero no en la línea n estará mal formada. Las reglas fonológicas se bloquearán cuando puedan crear una configuración así.

En su versión más intuitiva, esta restricción impone que para poder formar un pulso rítmico en cierto estrato (de la gradilla métrica) se haya formado también en todos los estratos inferiores.

Al hilo de la asignación del acento, digamos que el modelo teórico que hemos planteado para esta tesis nos parece especialmente fiel a la idea de columnas continuas. La proyección empieza desde el nivel segmental o melódico, y no se salta ni un solo paso.

En esto notamos una diferencia importante con los modelos de asignación del acento que se han hecho en Fonología Métrica que pasan directamente desde las ‘aspas’ métricas al acento, sin pasar por una formación de pies entre medias, que nosotros sí que consideramos.

Las consecuencias teóricas son sustanciales. Al tener en cuenta una formación de pies previa al acento, podemos prescindir de la **extrametricidad** que se ha postulado para los sustantivos en inglés. Por ejemplo, *baby* /^hbeɪbɪ/ no tendrá ya ninguna sílaba **extramétrica**, sino que el acento recae sobre la sílaba principal del troqueo, en una sucesión continua de proyecciones. Como el troqueo se forma antes de asignar el acento, no hay sílaba extramétrica. Así se resuelve un problema teórico, que simplemente deja de existir como tal. Y podemos trasladarlo al análisis del acento en latín y en otras lenguas, sin duda (véase el Apéndice III).

¹⁴ Un detalle que tenemos pendiente es saber si la sílaba tónica requiere peso en checo. Y si se les exige a las CS en particular. Su peso no podría venir de la propia CS, ya que tiene sólo una unidad de tiempo, pero Scheer habla de TVRT, asumiendo que están siempre en sílabas cerradas, que son pesadas por definición.

¹⁵ Scheer (2009: 423) cita dos ejemplos de líquidas sin vocal a principio de palabra, pero las considera como consonantes atrapadas. Se trata de *rdít se* (< *rdieti*) ‘ponerse rojo’ y *lhát* (< *lháti*) ‘mentir’.

Scheer considera sorprendente la ausencia de CS en posición de inicio absoluto, y la atribuye a que las CS (al igual que las consonantes atrapadas) se forman en un inicio ramificado, y por lo tanto no es una cuestión de contexto izquierdo *per se*. Las estructuras africadas que propusimos nosotros (recordemos el capítulo 6) pueden remedar este efecto.

¹⁶ Hemos encontrado una pequeña excepción a esta tendencia: *bɪnkout* ‘apestar’ (en Scheer 2009: 412) tiene una nasal en su contexto derecho inmediato. Eso sí, la nasal va seguida por una oclusiva.

¹⁷ Abundando en el paralelismo entre las CS checas y las CS atípicas, nos preguntamos si el sentido de la ramificación (más la aplicación de una metátesis) puede tener influencia en el comportamiento de estos dos tipos de CS con respecto al contexto vocálico / __V. Dejamos la pregunta planteada, y esperamos que se vaya resolviendo en el curso de estudios posteriores.

¹⁸ Consideramos que la regla se aplica también en los sufijos para formar pretéritos y participios, que vimos en las notas al capítulo 2. La hipótesis que barajamos ahora es que el sufijo /l/ es en realidad /-əl/ en su forma subyacente, y que después sigue el curso normal de la formación de CS en checo, por el cual la schwa no se filtra al plano silábico, sino que coalesce con la /l/.

¹⁹ Recomendamos Hale & Keyser (2000) para una explicación completa. Allí se explica que su modelo de referencia es la ‘incorporación’ de Baker (1988), por la cual una categoría gramatical puede formar un complejo con su objeto directo o con un adverbio sin perder su función sintáctica original.

²⁰ Este aspecto ya lo comentamos en el capítulo 6 (en la sección de notas finales), al que referimos para más detalles.

²¹ O al menos, no es una CS, sino una consonante atrapada, de acuerdo con la teoría de Scheer.

Apéndice I: La Fonología de la Rección

La **Fonología de la Rección** (FR) es una teoría desarrollada a finales de la década de 1980 y principios de los 90. Para este esbozo nos basaremos en Kaye (1990); Kaye, Löwenstamm & Vergnaud (1990); y Harris (1994).

La Fonología de la Rección nació en respuesta a la **fonología de reglas** propia de la **Fonología Generativa clásica**, en un intento de evitar el carácter arbitrario de las reglas del tipo $A \rightarrow B / _C$. Y más aún, se pretendía alinear la fonología con los avances en sintaxis que aportó el modelo de Principios y Parámetros¹. En un enfoque que al decir de sus autores no es nuevo, se hace una fonología con múltiples niveles de representación que permite pasar de un enfoque más **paradigmático** a otro más **sintagmático**.

Si fuera cierto que sintaxis y fonología se atienen, aunque sea en parte, a los mismos principios (1990: 194), la idea sería hacer toda una sintaxis de la fonología, mucho más allá de las coincidencias anecdóticas entre ambas disciplinas. El propósito es sacar a la luz mecanismos universales de la mente humana, tales como las asimetrías propias del lenguaje. Esta cuestión es de gran importancia: no se considera a la fonología como separada del resto del componente lingüístico, igual que la **forma lógica** (FL) tampoco se ha

tomado nunca así (KLV 1990: 194). Hay lugar para las diferencias entre sintaxis y fonología (el carácter recursivo sólo está en la sintaxis), pero no es descabellado concebir que algunos de los principios que están operativos en la sintaxis lo estén también en la fonología y viceversa (KLV 1990: 193).

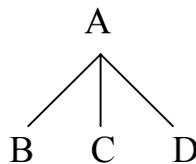
La FR asume tres reglas básicas. Una es la privatividad, que evita las **reglas de redundancia** para rellenar lo que falte, y que hace todos los procesos de armonía en positivo, sobre lo que está presente. Son univalentes, y no se puede expandir nada que no estuviera ya ahí. Parafraseando a Trubetzkoy, se diría que las **oposiciones privativas** nunca se convierten en **equipolentes** en el curso de la derivación. La segunda es la universalidad, donde el conjunto de procesos fonológicos se comporta como una **función** que proyecta (*map*) las representaciones iniciales a las finales. Un mismo objeto fonológico recibirá la misma **interpretación** a lo largo de todos los sistemas fonológicos, y las convenciones de ‘marcadez’ (*markedness*) son universales. La tercera es la no-arbitrariedad, que asume una relación directa entre los procesos fonológicos y su contexto. Se pone como ejemplo una expansión de tonos, que puede fundamentarse muy bien con reglas autosegmentales, y que en cambio resultaría **arbitraria** con un tratamiento de reglas (del tipo $A \rightarrow B / _ _ C$). (KLV 1990: 194-95).

La teoría debe someterse al peso de la evidencia empírica, desde luego. Un caso concreto es cómo explicar el acortamiento de vocales en sílaba cerrada que se observa en turco, en yawelmani, en inglés, en francés y en árabe, así como en wolof. Por ejemplo, en turco, *merak*, *mera:ki*, *meraktan* (‘ley’ en NOM, POSS y ABL).² Se propuso que los constituyentes silábicos entablaran relaciones sintácticas, y más

específicamente, de **rección**. Hay una **cabeza** que rige a los demás constituyentes de su **ámbito**, con relación de **mando-C**³. Las cabezas no necesitan ser regidas, pero ellas sí que rigen. La rección observa tanto la **adyacencia estricta** como la **direccionalidad estricta** (que a diferencia de la sintaxis no es paramétrica, sino universal; dentro de un constituyente, va siempre de izquierda a derecha).

Nótese cómo la adyacencia estricta y la direccionalidad sólo permiten que haya **estructuras binarias** (**Teorema de la Binariedad**). En efecto, queda prohibida cualquier **estructura ternaria** del tipo:

(1)



porque si B rige a los otros dos segmentos, se incumple la adyacencia estricta (entre B y D); y si C rige tanto a B como a D, se incumple la direccionalidad.

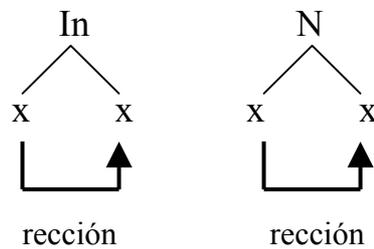
Una característica más de la FR es el **Principio de Proyección**, por el cual las operaciones siempre conservan la estructura original. Por consiguiente, queda descartado el resilabeo, y los rótulos de los nodos no varían a lo largo de toda la derivación.

Otra característica de la FR es que rechaza la existencia de un nodo **sílaba**. Su argumento es que, en contra de lo que suele asumirse, esta noción no tiene estatus preteórico⁴. Descartado el término ‘sílaba’, la FR delimita los nodos estructurales según las relaciones fonotácticas

que se observan en su interior. No todos los elementos se relacionan por igual entre sí: existen asimetrías y están sistemáticamente restringidas a determinados contextos, que se llaman **ámbitos fonotácticos** (*phonotactic domains*), y que se definen en términos de estructura de constituyentes.

La independencia fonotáctica entre inicio y núcleo justifica hacer una primera partición entre ‘inicio’ y ‘rima’. Sólo hay dos cabezas, **núcleo** (N) e **inicio** (In), donde sólo N es obligatorio. Cada cabeza delimita un ámbito de rección con el complemento que toma. Dentro de un nodo, la rección va de izquierda a derecha, por lo que en los nodos ramificados, sean inicios o núcleos, la cabeza siempre está a la izquierda:

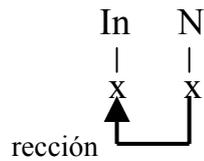
(2)



Estas estructuras representan el inicio ramificado y el núcleo ramificado. Obsérvese que se respeta el teorema de la binariedad.

Ha quedado establecido que dentro de un constituyente la rección va siempre de izquierda a derecha. Pero entre constituyentes el sentido de la rección se invierte: va de derecha a izquierda. Por eso los núcleos rigen a los inicios, y solamente los núcleos son obligatorios:

(3)



y también por eso es por lo que no puede haber dos inicios seguidos, aunque sí dos núcleos seguidos.

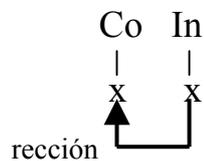
Nos faltaba un nodo silábico por analizar: la **coda** (Co). La FR no la considera como cabeza, sino como un simple adjunto al núcleo. Núcleo y coda forman el constituyente de la **rima** (R):

(4)



Dado que el N no rige a la coda, esta necesita de algún constituyente que la legitime. Lo hará un inicio que esté a su derecha, merced a que la rección entre constituyentes va de derecha a izquierda:

(5)



Hay tres ámbitos fonotácticos (*phonotactic domains*), que se definen según su estructura de constituyentes, por tanto:

- dentro de un inicio (/pl-/ en *play*).
- dentro de un núcleo (/ow-/ en *go*).
- dentro de un cúmulo [coda + inicio] (/nt-/ en *winter*).

Y en el plano de la estructura nos encontramos con una sucesión de inicios y núcleos unidos al esqueleto autosegmental, con las codas que ocasionalmente puedan aparecer. Como el núcleo es obligatorio, mientras que el inicio no lo es, puede haber dos N seguidos, aunque es una opción que tiende a evitarse. Sin embargo, jamás podemos encontrar dos inicios seguidos, como hemos dicho. Eso sí, se observa el principio de **maximización de inicios**, que eleva la cantidad de segmentos iniciales de sílaba hasta el máximo que permitan las condiciones que cada lengua impone sobre la estructura silábica. Por ejemplo, *restrain* contiene el inicio [str] y no las codas [st] o [s].

Una consecuencia interesante de la manera de legitimar las codas (mediante la rección que les da el inicio siguiente) es que no hay codas finales (desde Kaye 1990). ¿Qué ocurre con las consonantes finales? Pues se asignan a un inicio, pero este a su vez debe ir seguido de un núcleo que lo legitime, y la FR dice que ese N está vacío:

(6)

In	N
x	<u>x</u>

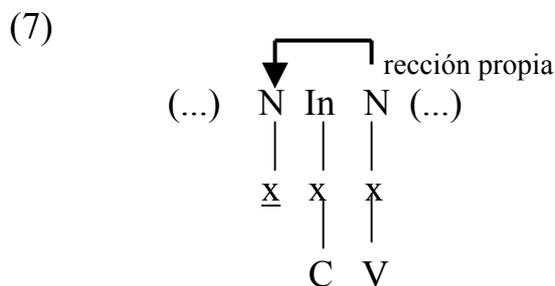
Inicio final de palabra:
 el núcleo vacío se indica con el punto esquelético subrayado.

Estos núcleos vacíos finales se licencian mediante un parámetro que decide si una lengua puede tener consonantes finales o no. Por ejemplo, el inglés puede tenerlas, pero el hawaiano no puede.

Los **núcleos vacíos** son una contribución característica de la FR. Para justificarlos, la FR toma como precedente la existencia de inicios

vacíos. De hecho, se postula que puede haber núcleos vacíos también en interior de palabra, y se afirma que gracias a ellos damos cuenta de las síncofes y alternancias entre vocal y ‘cero’.

Pero es imprescindible evitar la proliferación indiscriminada de categorías vacías. Por analogía con el **Principio de Categorías Vacías** de la sintaxis (PCV, o bien ECP en sus siglas inglesas), los núcleos vacíos deben estar regidos por una cabeza llena. En esto consiste la **Rección Propia**, que restringe la aparición de N vacíos. Para que un N pueda estar vacío en interior de palabra, debe recibir rección de otro N que esté lleno⁵. Como la rección entre constituyentes va de derecha a izquierda, la rección propia sigue esa misma dirección:



Rección de núcleos vacíos internos de palabra. C y V son vocales y consonantes cualesquiera. El aspa subrayada indica que está vacía, sin segmento asociado.

La necesidad de rección propia se apoya en el hecho cierto de que una vocal se puede suprimir sólo si está adyacente a una vocal que no se suprima. Además, en todos los casos conocidos, esa vocal está a la derecha del núcleo vacío⁶.

Concluiremos esta introducción con una vuelta al plano melódico, con el fin de estudiar con más detalle la composición de los segmentos fonológicos. Los primitivos de la FR no son los **rasgos**, sino los

elementos. Su número es muy pequeño (en torno a diez solamente), porque la FR sigue la orientación minimista y prefiere pecar por defecto, en su afán de evitar la sobregeneración⁷. A diferencia de los rasgos, los elementos de la FR tienen la opción de existir por sí mismos o bien de formar compuestos o complejos con otros elementos, como hacen los elementos de la materia física.

Una segunda diferencia es que los elementos tienen independencia y plena especificación. Por eso pueden ser interpretados fonéticamente en cualquier momento, a lo largo de toda la derivación. En virtud de esto, la derivación fonológica no se convierte en una operación que vaya transformando los objetos fonológicos abstractos en objetos físicos cada vez más concretos. En la FR, los procesos proyectan los objetos fonológicos a otros objetos fonológicos, más que a objetos fonéticos. La derivación, por tanto, es una función estrictamente generativa que define un conjunto de representaciones fonológicas bien formadas, la gramaticalidad de las sartas fonológicas.

Se evita así la sobregeneración debida a la infraespecificación y las reglas de redundancia, porque ya no hay necesidad de completar los valores no-marcados sin especificar.

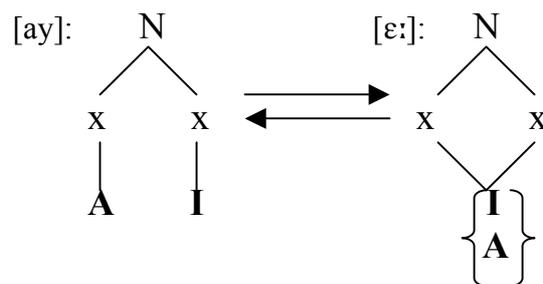
La tercera diferencia con los rasgos distintivos está en que los elementos son todos **monovalentes**: simplemente se tienen o no tienen; entran o no a formar parte de la composición de un segmento, pero no hay especificación positiva o negativa.

Para dar una idea de las representaciones basadas en elementos, vamos a ver cómo se caracterizan las vocales. Los elementos fundamentales de las vocales son **A**, **I**, y **U** (nótese que los elementos

se representan siempre con negrita). La interpretación fonética de **A**, **I**, y **U** corresponde a las **vocales cardinales** respectivas: esto es, **A** se interpreta como [a], **I** como [i] y **U** como [u]. Partiendo de esta base, se pueden combinar los elementos para definir otras vocales. Por ejemplo, [e] es [**A,I**], y [o] corresponde a [**A,U**].

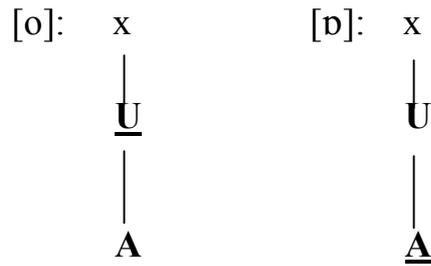
Cuando los elementos están distribuidos en dos unidades de tiempo, tenemos un diptongo, y se puede pasar de diptongo a monoptongo y viceversa con sólo cambiar las líneas de asociación de los elementos al esqueleto:

(8)



Pero aún hay que refinar un poco el modelo para dar cabida a todas las vocales de un sistema tan complejo como el inglés. Para conseguirlo hay dos mecanismos: uno es establecer relaciones de dependencia entre los elementos, de forma que si hay más de un elemento en un segmento, uno de los elementos sea la cabeza y los demás sean dependientes. Por ejemplo, para diferenciar entre [o] y [ɒ], basta con invertir la relación:

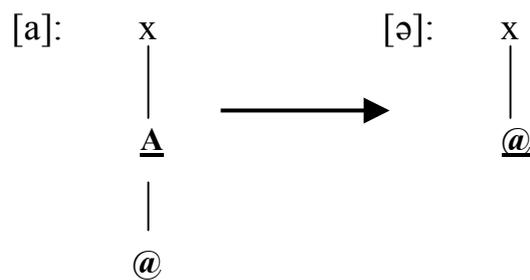
(9)



N.B.: La convención gráfica es subrayar el elemento que actúa como cabeza.

El otro mecanismo consiste en postular un elemento adicional, que se llama @ y cuya realización fonética es la de la vocal neutra **schwa**, que se representa como [ə]. Gracias al elemento @ podemos dar cuenta de las vocales [ɪ] y [ʊ], que son [@,I] y [@,U], respectivamente. Además, asumiendo que el elemento neutro está presente en todas las representaciones, podemos simbolizar la reducción vocálica del inglés en forma de pérdida de elementos. En concreto, la reducción de [a] a [ə] es como sigue:

(10)



Otra utilidad de estas representaciones con elementos y esqueleto autosegmental es la posibilidad de dar cuenta de las distintas pronunciaciones de una misma vocal en dialectos diferentes: para conseguirlo basta con cambiar la asociación entre elementos y gradilla esquelética.

La @ sirve de paso para dar cuenta de las posiciones nucleares ‘vacías’. La FR dice no están en absoluto vacías, sino que contienen un material latente en forma de @, junto con el cual forma el ámbito potencial de legitimación autosegmental. La activación de su potencial legitimador depende de que se cumplan ciertas condiciones. Cuando la posición recibe rección propia, puede quedar como N vacío, y cuando no, debe expresarse fonéticamente el material latente.

En lo que respecta a las consonantes, además de los elementos **A**, **I**, **U** y @, hay algunos elementos más que entran en su constitución y que pasamos a ver someramente. En relación con el lugar de articulación, ? es el elemento de estrechamiento prolongado del tracto vocal a la altura de los labios, **h** se define como “ruido” y es un flujo de aire turbulento que se observa en la aspiración, y **R** es el elemento coronal. **N** es el elemento nasal. En las cuerdas vocales, **L** es el elemento laxo y **H** es el tenso; el elemento laxo está en las obstruyentes sonoras, mientras que el tenso está en las insonoras o *fortis*⁸. Para obtener una idea más precisa de cómo se definen los segmentos consonánticos a partir de los elementos, veamos algunos ejemplos:

(11)

p = [U, <u>?</u> , h , H , @]	t = [R , ?, h , H , @]	d = [R , ?, h , @]
m = [U, ?, N , @]	ŋ = [?, N , @]	h = [h , @]
f = [U, <u>h</u> , H , @]	s = [R , <u>h</u> , H , @]	ʃ = [I , R , <u>h</u> , H , @]
l = [R , <u>?</u> , @]	ɹ = [R , @]	tʃ = [I , R , ?, H , @]

N.B.: El símbolo ‘ɹ’ representa la ‘r’ típica del inglés. Los ejemplos de este diagrama provienen de Toft (2002: 129), que hace suyo el criterio de que el elemento @ está presente en todos los casos.

¹ El modelo de **Principios y Parámetros** arranca de los trabajos de Chomsky en 1981 (véase Haegeman 1991, 1994)). Su planteamiento es que existe una **Gramática Universal** innata formada por una serie de principios universales, a los que se da especificidad mediante los correspondientes parámetros, que son propios de cada lengua. El modelo tiene dos etapas. La **Teoría de Rección y Ligamiento** (*Government and Binding*) está muy bien explicada en el trabajo de Haegeman, y también en Radford (1988), que se centra en la lengua inglesa. Para la otra etapa, la del **Programa Minimista** o **Minimalista**, se pueden consultar Chomsky (1995), Radford (2004), Lasnik & Uriagereka (2005), o Bošković & Lasnik (2007).

² Dicho de otro modo, estas lenguas parecen prohibir las vocales largas en sílabas cerradas.

³ El ‘mando-C’ o ‘mando de constituyentes’ es una noción de la Teoría de Principios y Parámetros que define una relación estructural entre constituyentes. En términos intuitivos, alude a los nodos hermanos y a toda la línea de “sobrinos” de cualquier orden. En términos formales se define como:

Mando-C:

Un nodo A manda-c a un nodo B si y sólo si:

- (i) A no domina a B y B no domina a A; y
- (ii) El primer nodo ramificado que domina a A también domina a B

Según la definición original de Reinhart (ver Haegeman 1994: 134), que Chomsky vuelve a formular en *Barriers* (ver 1994: 137) como:

Mando-C:

A manda-c a B si y sólo si A no domina a B y todo X que domine a A también domina a B.

⁴ Según Harris (1994: 45), en las sociedades alfabetizadas la “sílabas” se deriva de las convenciones que rigen la escritura y la versificación (y que varían a lo largo de las distintas tradiciones ortográficas). Después, el término se suele enseñar en los colegios como un medio para etiquetar algún aspecto de la realidad fonológica, pero luego no está nada claro cuál es exactamente esa realidad.

Pero en nuestra opinión, mantener un nodo σ tiene ventajas como la de explicar la rección entre núcleos (nodos N), y la de la posibilidad de proyectar a niveles más altos, en concreto, a pies métricos, lo cual es difícil de hacer sólo con O y N aislados.

⁵ De hecho, se ha propuesto que las dos estén sujetas a la misma restricción universal, que es el **Principio de Categorías Vacías** (PCV, o también ECP en sus siglas inglesas), una de cuyas subcláusulas pide que una categoría o posición siga regida por su cabeza, incluso cuando haya quedado vacía o vacante a raíz de un movimiento sintáctico o de la disociación (*delinking*) de material melódico. Podríamos decir que un N vacío recibe rección propia de su N de cabeza en la misma medida en la que, por ejemplo, la huella del complemento de un SN (sintagma nominal) que ha sufrido un movimiento Q (*WH-Movement*) está regida por su cabeza verbal.

⁶ El requisito de adyacencia sugiere que la vocal diana (*target*) y la desencadenante de una síncope entran en una relación de legitimación (*licensing*) en algún nivel de la proyección nuclear. Por ejemplo, en *family* se puede elidir la vocal del N₂ gracias a que recibe rección propia de la del N₃. El hecho de que en unas lenguas la síncope sea obligatoria, en otras sea opcional, y en otras sea imposible, indica que la rección propia no se impone estrictamente en todos los sistemas.

La rección propia asegura que una vocal sincopada esté siempre adyacente a una vocal sin sincopar. Esta restricción deriva por consiguiente el patrón de alternancia VC0CV0C0... que es característico de las lenguas donde la síncope se repite o itera a lo largo de toda la palabra. En cambio, se evita una secuencia del tipo C0C0CV..., ya que el N₂, aunque recibe rección del N₃, es a su vez incapaz de dar rección propia al N₁. De modo que un efecto de la restricción que impone la Rección Propia es excluir la adyacencia de N vacíos en las representaciones fonológicas.

Aún podemos llevar más lejos el paralelismo y decir que no se pueden establecer interacciones sistemáticas entre posiciones vecinas cuando se interponga una posición vacía. Al igual que podemos contraer *She's [the woman]_i I want to (wanna) talk to [0]_i*, pero no en *She's [the woman]_i I want [0]_i to (*wanna) be president*, podemos pensar en cómo no se manifiestan dependencias fonotácticas entre inicios cuando entre ellos se interpone un N vacío.

⁷ Sólo cuando el modelo parezca infragenerar se plantea como problema para que la investigación lo intente resolver.

⁸ Toft (2002:128) da una descripción precisa de cada elemento, basándose en Harris (1994, 1995). Según recoge esta autora, **A** es la masa de energía central, por convergencia de F_1 y F_2 , y se encuentra en las uvulares y faríngeas. **I** es un F_1 bajo acoplado a un pico espectral alto (convergencia de F_2 y F_3), y está en las consonantes palatales y palatalizadas. **U** es un pico espectral bajo (convergencia de F_1 y F_2) que está en las vocales y consonantes labiales. **@** es una estructura de formante disperso, esto es, sin ningún pico espectral que sobresalga, y forma una base sobre la que se superponen las otras resonancias. Está en las consonantes velares. **R** es un conjunto de transiciones de formantes asociadas a las coronales. **h** es una energía aperiódica, que resulta de una constricción estrecha que produce un flujo turbulento del aire. **ʔ** es un descenso abrupto y sostenido de la amplitud general, y que cuando se manifiesta independientemente es una oclusiva glótica. **N** es el murmullo de una banda ancha de baja frecuencia. **H** es una frecuencia fundamental alta, con las cuerdas vocales tensas o rígidas. Y **L** es una frecuencia fundamental baja, con las cuerdas vocales laxas.



Apéndice II: La derivación fonológica

En este apéndice vamos a hacer una derivación fonológica siguiendo las pautas del modelo teórico que hemos propuesto. Con ello pretendemos dar una visión de conjunto sencilla que facilite la lectura de toda la tesis. No citaremos referencias porque ya están en el cuerpo del documento y este apartado sólo pretende ser un repaso de lo ya expuesto.

Vamos a empezar por abajo, por las unidades más pequeñas, que son los rasgos infrasegmentales, y vamos a ir ascendiendo paso a paso por toda la maquinaria fonológica hasta llegar al acento de palabra. Por el camino vamos a recorrer los segmentos, las unidades de tiempo del esqueleto autosegmental, los nodos silábicos, las moras o unidades del esqueleto moraico, y finalmente, los pies métricos. Como muestra usaremos la palabra *seven*, que pertenece al inglés cotidiano y que es susceptible de formar una CS. Tengamos en cuenta que la derivación va a quedar “boca abajo”, puesto que empezamos por lo más básico pero el documento se desarrolla en sentido descendente.

En primer lugar tenemos los rasgos autosegmentales. Recordemos que se pueden hacer representaciones a base de elementos (como hace la Fonología de la Rección), pero que nosotros nos atenemos a los rasgos. Estamos en el plano melódico, y los rasgos tienen su

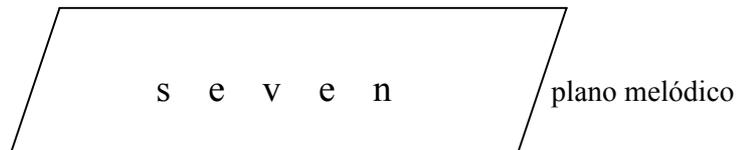
geometría propia, con disposición en planos autónomos, de acuerdo con el modelo autosegmental (por eso cada rasgo tiene su propia matriz). Vamos a ver sólo los rasgos de los dos primeros segmentos, para mayor claridad:

(1)

s	e	v	e	n
[+conso]	[-conso]			
[-sonante]	[+sonante]			
[+contin]	[-alto]			
[-sonoro]	[-bajo]			
[+estrid]	[-retraído] etc.			

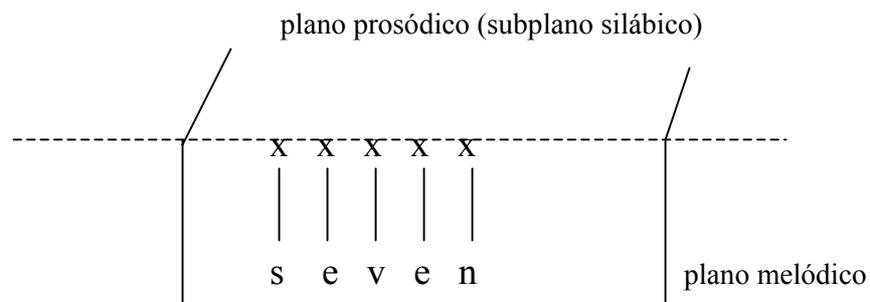
Los rasgos de cada segmento confluyen en un nodo raíz, y en el plano melódico se forma una secuencia con todos los segmentos alineados:

(2)



Ahora los segmentos se van a asociar a las unidades de tiempo del esqueleto inferior, que es el que articula el plano melódico con el plano prosódico:

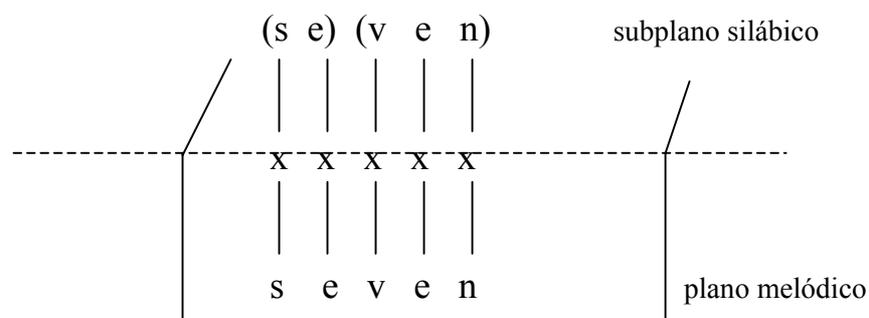
(3)



En este ejemplo vemos que todas las asociaciones son 1:1, esto es, cada segmento se asocia con un aspa del esqueleto. Pero recordemos que existen también asociaciones 2:1 (en las africadas) y asociaciones 1:2 (en vocales largas y en consonantes geminadas).

Ahora se van a licenciar los segmentos mediante la asignación de nodos silábicos. Recordemos que se observan el Principio de Secuenciación Sonora (PSS) y el Principio de Maximizar los Inicios:

(4)



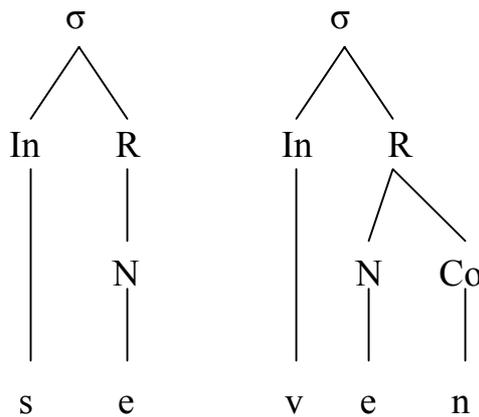
Hemos proyectado directamente los segmentos al siguiente plano, y los hemos agrupado en sílabas. Si lo vemos con más detalle, la estructura resultante es:

(5)

$$(\mathbf{s}_{In} (\mathbf{e}_N)_R)_\sigma \cdot (\mathbf{v}_{In} (\mathbf{e}_N \mathbf{n}_{Co})_R)_\sigma$$

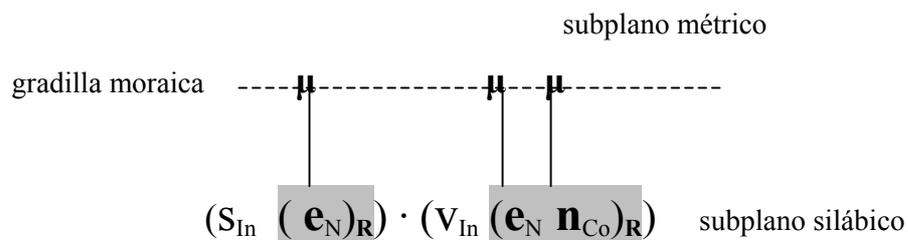
Esta es una representación con rótulos, que traducido a una estructura arbórea nos daría:

(6)



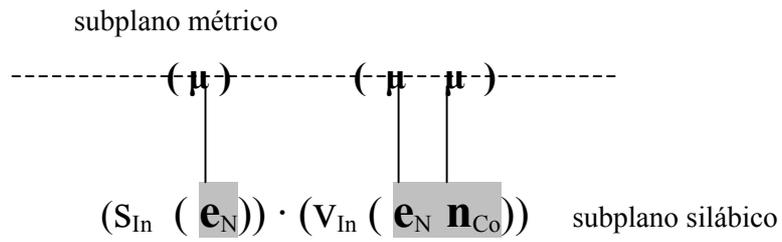
Ahora están todos los segmentos licenciados y estructurados. Es el momento de pasar al otro subplano prosódico, que es el métrico. Para ello tenemos otra interfaz que permite articularlos, que es el esqueleto superior o ‘gradilla moraica’. El mecanismo es igual que en el esqueleto inferior: asociar elementos del plano con los nudos de la interfaz:

(7)



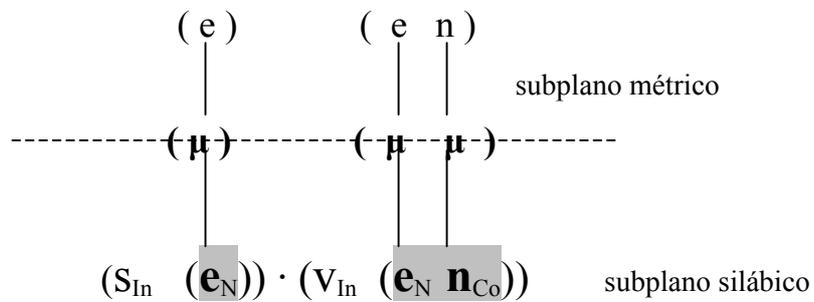
Vemos que se respeta el ‘peso por posición’: solamente los segmentos de la rima (núcleo y coda, marcados en negrita y resaltados en gris), se asocian a moras, pero no así los de los inicios. Los paréntesis de agrupación de segmentos en una misma sílaba se van a respetar en la traslación al plano siguiente:

(8)



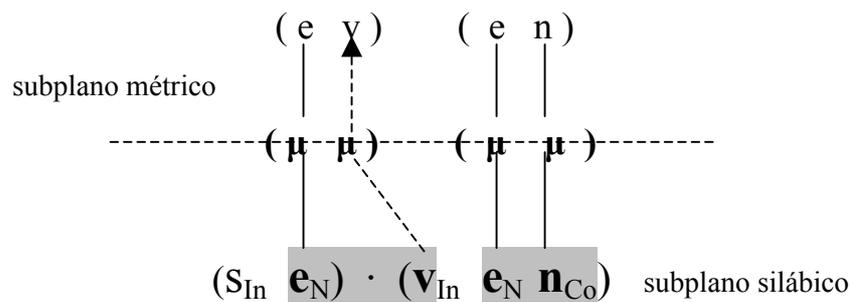
Esa misma agrupación se proyecta al subplano métrico:

(9)



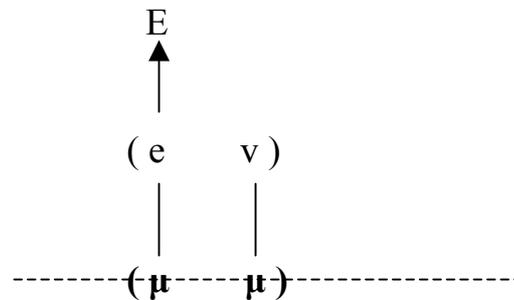
La primera sílaba es ligera, sólo ha proyectado una mora al siguiente nivel. Pero queremos formar un troqueo, y para eso hace falta una sílaba pesada en la parte izquierda. A tal efecto, hacemos una captura de mora, con el inicio de la sílaba siguiente:

(10)



Ahora sí que tenemos dos moras agrupadas en torno a la primera sílaba de *seven*, y por eso podemos proyectar una supermora al nivel siguiente. Expresado en estructura de frase escueta:

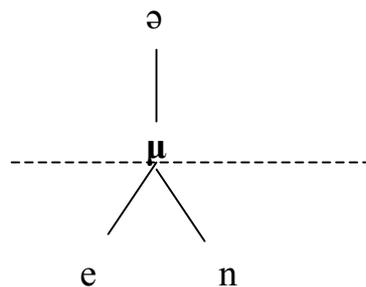
(11)



Por convención, hemos representado la proyección de la ‘e’ a supermora con su correspondiente mayúscula, una ‘E’.

Por su parte, la otra porción del troqueo es débil y queda como ‘resto’. Sólo se va a computar una unidad de peso, porque la proyección es semitransparente:

(12)



La vocal átona se ha neutralizado a schwa, como manda la fonología del inglés. En nuestro modelo, diremos que se han filtrado selectivamente algunos rasgos de esa vocal (por lo que otros no se han filtrado), en una proyección que consideramos semitransparente. En

cambio, la proyección de las vocales tónicas es transparente, por lo cual filtran todos sus rasgos al plano métrico, pasan intactas.

Con la suma de los dos componentes podemos formar el troqueo. El panorama completo ahora es:

$$(13) \quad \begin{array}{c} \left[\begin{array}{cc} E & \\ (e \ v) & \text{ə} \end{array} \right] \\ \text{-----}(\mu \ \mu)\text{-----}\mu\text{-----} \end{array}$$

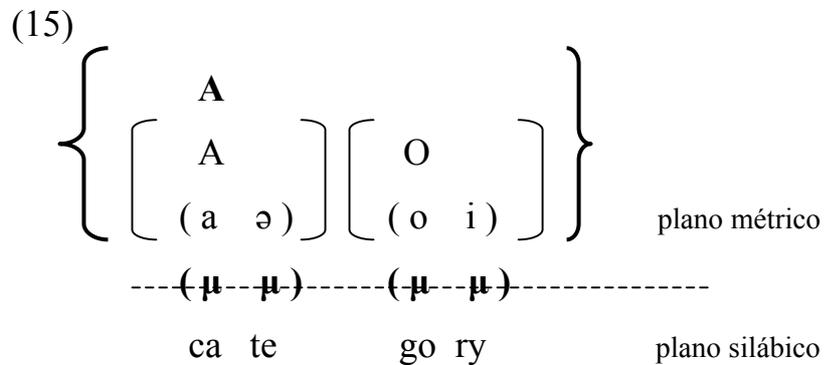
[Para facilitar la exposición, hemos omitido esta vez las líneas de asociación hacia el plano métrico que habíamos expresado en los diagramas previos. Tampoco expresamos ya los componentes del plano silábico, para poder centrarnos mejor en el plano métrico.]

La parte alta vuelve a proyectar, un nivel más arriba, para formar el acento de palabra, que representamos en negrita:

$$(14) \quad \begin{array}{c} \mathbf{E} \\ \left[\begin{array}{cc} E & \\ (e \ v) & \text{ə} \end{array} \right] \\ \text{-----}(\mu \ \mu)\text{-----}\mu\text{-----} \end{array}$$

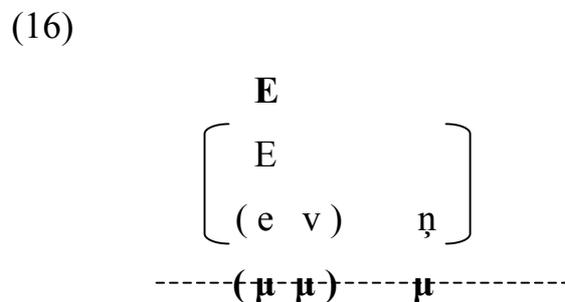
En este caso la proyección es vacua: la cabeza del troqueo es también la cabeza de la palabra, porque sólo hay un pie. Pero en una palabra como *category* hay dos pies, y entonces hay que elegir cuál de las supermoras proyecta un nivel más arriba. La que lo haga tendrá el

acento principal de palabra, y la que no lo haga quedará como acento secundario:



Las {llaves} marcan el conjunto de la palabra. Los [corchetes] marcan los dos troqueos. La **A** en negrita corresponde al acento de palabra. La **A** y la **O** mayúsculas son los acentos que corresponden a las supermoras. Como la **O** no proyecta al nivel más alto, queda como acento secundario. Observemos cómo el acento se obtiene a base de proyecciones sucesivas y continuas, que pasan de un nivel a otro sin saltarse ninguno.

Finalmente, para la proyección de una CS nos basta con cambiar la schwa de *seven* por la propia ‘n’, que pasa a ser silábica:



A tal fin se ha dado una confluencia (*conflation*) entre los dos segmentos, donde la 'n' expresa toda su matriz sonora, mientras la vocal queda inexpresada pero es la que aporta el nodo nuclear.

Recordemos que esta notación en estructura de frase escueta se corresponde con la de aspás y nodos de otros sistemas, y que no hay más que sustituir los rótulos para obtenerla, directamente:

(17)

$$\begin{array}{c}
 \mathbf{M} \\
 \left[\begin{array}{cc}
 \mathbf{M} & \\
 (\mu \mu) & \mathbf{x}
 \end{array} \right] \\
 \text{-----}(\mu \mu)\text{-----}\mu\text{-----}
 \end{array}$$

Con esto queda delineada una derivación completa de una palabra cualquiera del inglés.

Apéndice III: El acento en latín

Para explorar la capacidad predictiva del modelo vamos a hacer un estudio del acento en latín, aplicando el modelo teórico de esta tesis a una lengua distinta del inglés y a un fenómeno bien distinto de las consonantes silábicas. Y de paso, volvemos sobre el modelo para facilitar y mejorar su comprensión.

Todo el análisis que sigue está basado en conceptos ya expuestos en esta tesis, pero sobre todo, en la manera de organizar el espacio que propusimos en el capítulo 3. Partiendo de esta base, empezamos por la distinción entre cantidad y peso. Recordemos que **cantidad** se refiere al número de constituyentes segmentales de la sílaba, mientras que el **peso** se limita al número de constituyentes silábicos que se computan en el sector métrico. Estamos asumiendo el ‘peso por posición’ de Hayes (1995), donde los inicios no se cuentan. Decimos que el núcleo es la fuente primaria del peso¹, por ser cabeza de la sílaba, en tanto que la coda es (a lo sumo) coadyuvante².

Recordemos que las **moras** quedan por encima del nodo silábico, y que además funcionan como interfaz con el plano métrico, con el mismo sistema que el esqueleto de tiempo conectaba los planos segmental y silábico. Y que el paso del plano silábico al métrico está regido por el grado de **transparencia**; en función de cuánto se pueda

filtrar hemos dicho que las proyecciones pueden ser transparentes, semitransparentes y opacas (los inicios son opacos por definición).

Establecemos que el latín es una lengua sensible al peso, y diremos también que organiza el plano métrico a base de pies. En nuestro modelo todos los pies son binarios, por lo que descartamos los **dáctilos** (X x x). Para el latín vamos a descartar igualmente los **yambos** (x X), dada la inexistencia de palabras oxítonas en esta lengua. Nos quedamos entonces con la idea de que los pies latinos son todos troqueos, con una estructura (X x), o sea, binaria y con la cabeza a la izquierda.

Basándonos en Hayes (1995), vamos a distinguir dos tipos de troqueos (ambos con la cabeza a la izquierda como troqueos que son). Hayes habla de ‘troqueo moraico’ y ‘troqueo silábico’, pero nosotros los vamos a redefinir como ‘troqueo simple’ y ‘troqueo compuesto’. El **troqueo simple** se define así:

$$(1) \quad \left[\begin{array}{cc} M & \\ (\mu) & (\mu) \end{array} \right] \quad \begin{array}{l} \textit{l\i{nea} 1} \\ \textit{l\i{nea} 0} \end{array}$$

La M representa una **supermora**, que procede de la proyección de dos moras agrupadas. Los paréntesis propios que hemos asignado a cada mora (μ) indican que cada una viene de una sílaba distinta, pero las dos se agrupan en el troqueo, y de ahí la proyección a supermora.

El **troqueo compuesto** parte de una supermora y la vuelve a proyectar, en una estructura recursiva. Para formar este troqueo,

empezamos por una supermora, a base de dos moras agrupadas (vemos que sólo hay un paréntesis, porque las dos provienen de la misma sílaba):

$$(2) \quad \begin{array}{ll} M & \textit{línea 1} \\ (\mu \mu) & \textit{línea 0} \end{array}$$

Esta estructura es similar a la que tenía el troqueo simple, pero no tiene corchetes o **rótulo** de troqueo. Esto es, la supermora no se licencia todavía como pie, sino que toma un nuevo complemento, un nivel más arriba:

$$(3) \quad \begin{array}{ll} M^2 & \textit{línea 2} \\ (M \mu) & \textit{línea 1} \end{array}$$

y así es como se va a formar el pie³. En suma, la estructura del troqueo compuesto es:

$$(4) \quad \begin{array}{ll} \left[\begin{array}{ll} M^2 & \\ (M & \mu) \end{array} \right] & \textit{línea 2} \\ & \textit{línea 1} \\ (\mu \mu) \quad \mu & \textit{línea 0} \end{array}$$

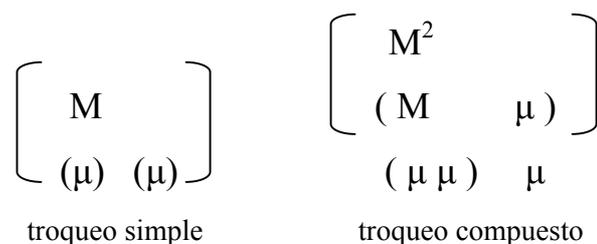
Por lo tanto, el troqueo simple se forma un nivel más abajo, y proyecta desde la línea de las moras, pero el troqueo compuesto se forma un nivel más arriba (es más complejo) y tiene que formar primero una

supermora para después formar la estructura $(M^2 (M \mu))$, que es la que da lugar al pie.

Los **corchetes** [x] indican que se ha formado un pie, trocaico en este caso. La falta de corchetes en la primera proyección de la supermora $M (\mu \mu)$ del troqueo compuesto es lo que marca la diferencia entre supermora propiamente dicha y troqueo simple. También la procedencia de las moras, según sean tautosilábicas o heterosilábicas. Pero por lo demás, el troqueo compuesto es una proyección iterada de la estructura básica $(X (X x))$, esto es, binaria y con la cabeza a la izquierda.

Casi podríamos decir que el troqueo compuesto consta de dos pies incrustados, un troqueo simple encima de otro. Pero no es del todo exacto. En la línea de base hay una supermora, que no llega a licenciarse como troqueo simple porque no se **licencia** como pie (por eso no tiene los corchetes). Es importante ver en qué momento se licencia un pie como tal. El troqueo simple se licencia nada más formarse la supermora, pero en el troqueo compuesto se necesita iterar el mecanismo de proyección $X (x x)$ para licenciar el pie. Gráficamente:

(5)



Un último detalle en lo referente a la estructura. Es de suma importancia la distribución asimétrica del material. La diferencia entre **cabeza** y **complemento** es determinante. El complemento sólo es coadyuvante, y como tal, sólo se limita a ayudar a la cabeza, por lo cual tiene una proyección semitransparente. Todos los complementos son **semitransparentes**, por definición.

Es hora de pasar al análisis concreto. Establecemos una primera diferencia entre verbos y nombres. El latín decide (de manera paramétrica) sentar una distinción categorial entre nombres y verbos (que no se sigue en otras lenguas, como el francés, el húngaro o el polaco). Los verbos forman troqueos simples, y los nombres forman troqueos compuestos. Eso sí, todo se atiene al patrón trocaico, como ya dijimos desde el principio.

Vamos a empezar por los verbos. Decimos que siempre forman un troqueo simple. A tal efecto, la sílaba final queda como complemento y la penúltima recibe el acento. Siempre es así. Da igual *vulnero* ‘herir’ que *amabo* ‘amaré’ o *desponso* ‘desposar’. La penúltima sílaba aporta una mora, y la sílaba final aporta la otra mora. Eso es todo. Por ejemplo:

(6)

$$\left[\begin{array}{cc} M & \\ (\mu) & (\mu) \end{array} \right] \text{troqueo simple}$$

vul ne ro

N. B.: Para facilitar la exposición en lo sucesivo, vamos a numerar las sílabas empezando por el final, con lo que una palabra trisílaba está formada por $[\sigma_3 - \sigma_2 - \sigma_1]$.

Por eso el acento no es sensible al “peso” de la σ_2 , que recibe el acento en una palabra trisílaba como *vulnero* pese a ser una sílaba “ligera”.⁴ Tenga el “peso” (la cantidad) que tenga, la σ_2 es el que recibe el acento de palabra por ser la cabeza de un troqueo simple, donde la σ_1 hace el papel de complemento. Es de resaltar que la σ_1 no es **extramétrica** en modo alguno, sino que forma el complemento del troqueo.

Al ser el complemento o ‘resto’, la proyección de σ_1 es semitransparente, por lo que el material de más no afecta al peso. Se limita a complementar a la σ_2 , y al ser coadyuvante, sólo puede aportar una mora, que es ‘cantidad suficiente’. El resto de la cantidad no se computa en forma de peso, por eso da igual que la σ_1 sea “ligera” (*vulnéro*) que “pesada” (*vulnéras*, *vulnérat*).

También debido a que se forma el troqueo simple, la mora de la izquierda (la de la σ_2) no puede alcanzar mayor complejidad, y por eso el acento permanece siempre allí. Por eso tenemos *vulnerámus*, *vulnerábit*, con el acento fijado siempre en σ_2 :

(7)

$$\left[\begin{array}{cc} \text{M} & \\ (\mu) & (\mu) \end{array} \right] \text{troqueo simple}$$

vul ne ra mus

Llega el turno de los nombres. El **prototipo** de nombre tiene un troqueo compuesto. Partimos de un nombre como *mensa* ‘mesa’, que es perfecto porque sólo tiene dos sílabas (no hay mucho que rebuscar a la hora de asignar el acento), y la σ_2 además es pesada (sin comillas esta vez, porque tiene dos elementos en la rima que proyectan dos moras). La estructura resultante es:

$$(8) \quad \left[\begin{array}{cc} M^2 & \\ (M & \mu) \end{array} \right] \text{troqueo compuesto}$$

$$\begin{array}{cc} (\mu \mu) & \mu \\ | & | \\ m e n & sa \end{array}$$

Y todo lo demás son variaciones sobre este prototipo. Vamos a probar con otras palabras. Empezamos por las palabras de menor cantidad, aquellas que no alcanzan para formar un troqueo compuesto. El primer escalón es una palabra de dos sílabas pero cuya σ_2 es “ligera” (no tiene cantidad suficiente para formar una supermora). Por ejemplo, *rosa* ‘rosa’, que debe conformarse con un troqueo simple:

$$(9) \quad \left[\begin{array}{cc} M & \\ (\mu) & (\mu) \end{array} \right] \text{troqueo simple}$$

$$\begin{array}{cc} ro & sa \end{array}$$

porque no tiene material suficiente para iterar la estructura y llegar al troqueo compuesto. Pero se admite porque al menos es capaz de

proyectar un troqueo. Es de señalar que *domus* ‘casa’ también forma un troqueo simple. La σ_1 no cuenta nunca como pesada, porque al ser el complemento, su proyección es semitransparente, y por tanto sólo se computa una mora.

Pasamos a los monosílabos, que tienen menos material todavía. Analizamos *vir* ‘varón’, y vemos que ahora se necesita computar la coda para sumar el peso. La estructura es:

(10)

$$\left[\begin{array}{c} M \\ (\mu \quad \mu) \end{array} \right] \text{troqueo reducido}$$

v i r

La estructura del **troqueo reducido** es como la del troqueo simple⁵, pero a diferencia del de *rosa* (o del de *vulnero*), sólo consta de una sílaba, y por eso las dos moras están agrupadas en un solo paréntesis. Y es que hay más transparencia: a falta de otros núcleos (o sílabas) que computar, el mecanismo de transparencia se ha abierto para que entren en juego como coadyuvantes los segmentos de la coda. Se ha hecho totalmente sensible al peso esta sílaba, que es sílaba única. Eso sí, no se permiten monosílabos ligeros, porque sería imposible formar un troqueo con ellos; no habría de dónde sacar material. En efecto, el latín veta las palabras léxicas (*content words*) de sólo una mora.

Cuando había material de menos, nos hemos movido desde la estructura prototípica de *mensa* (el troqueo compuesto, formado por la supermora y su complemento) hacia abajo. Ese troqueo compuesto se formaba en la línea 2 de la estructura métrica. Cuando hay menos

material, bajamos un nivel, a ver si basta: es lo que hacemos en *rosa* o *domus*, abriendo el mecanismo de licenciar a un simple troqueo simple. Si no podemos licenciar un troqueo compuesto, entonces descendemos en la estructura y licenciamos un troqueo simple.

Y si ni siquiera tenemos un núcleo (o sílaba) que sirva de complemento del troqueo simple, obtenemos el material de la propia sílaba, esto es, de la coda que la complementa. Formamos un troqueo reducido. En realidad, *vir* es idéntica a la supermora del troqueo compuesto de *ménsa*. Se ha abierto la transparencia (al descender en la estructura) para computar material de la rima, de la coda. Y una vez más, como el complemento es por definición semitransparente, la coda compleja de *rex* ‘rey’ aporta el mismo peso que la de *vir*.

Un aspecto importante del esquema es que no hay elementos extramétricos en ningún momento. La σ_1 no es en modo alguno extramétrica, sino que es el complemento necesario del troqueo, y por tanto se comporta como tal, no se excluye del recuento.

Otro aspecto a tener en cuenta es la semitransparencia de los complementos, ya referida, que explica por qué *rosa* y *domus* tienen el mismo patrón acentual (la σ_1 sólo puede computarse como ligera, por su condición de ‘resto’), y por qué *vir* y *rex* sólo tienen dos moras (de nuevo, la coda sólo puede aportar una mora, por ser el complemento)⁶.

Un tercer detalle es que para formar el troqueo se toma como referencia el nivel de los núcleos. Para proyectar moras en el troqueo canónico hace falta la presencia de dos núcleos. Por eso *domus* es *dómus* y no *domús*.⁷ Compárese:

(11)

$$\begin{array}{cc} \left[\begin{array}{cc} \text{o} & \\ \text{o} & \text{u} \end{array} \right] & \left[\begin{array}{cc} \text{u} & \\ \text{u} & \text{s} \end{array} \right] \\ (\mu) (\mu) & (\mu \mu) \\ \text{do mus} & \text{do m u s} \\ \text{dómus} & *domús \end{array}$$

Desde aquí volvemos al troqueo compuesto. En realidad, es una combinación de los dos mecanismos que hemos visto en las palabras más cortas. La base es un troqueo simple como el de *rosa*, donde se computan los núcleos, y donde la σ_1 hace de complemento y como tal es semitransparente.

Pero sobre esta base se añade un nuevo nivel de complejidad. Se itera la proyección, y ahora la cabeza es sensible al peso (a diferencia del troqueo simple, cuya cabeza es monomoraica)⁸. En la medida de lo posible, el troqueo compuesto forma la supermora con el propio material de la σ_2 (obtiene el peso de ahí), exactamente igual que se hacía en el troqueo reducido con *vir* o con *rex*. Esto es, se abre la transparencia en esa ‘zona caliente’ que es la σ_2 , y por eso se computan todos los elementos de su rima. Es una proyección máximamente transparente de la σ_2 . Podemos tomar el material del núcleo cuando es ramificado (que aporta más cantidad, como en *ami:cus* ‘amigo’), o bien tomarlo de la coda (como en *argentum* ‘plata’). Nótese de nuevo que la coda es coadyuvante, y no podrá aportar más de un elemento al peso. Y por eso *Calixtus* tiene la σ_2 igual de pesada que *argentum*, porque aunque tenga más cantidad, tiene el mismo peso.

Estamos siguiendo un poco el camino inverso: antes descendimos primero desde *mensa* a una estructura más simple (*rosa*), y después a otra más simple todavía (*vir*). Porque los troqueos se iban formando en niveles cada vez más bajos:

(12)

TROQUEO COMPUESTO

$$\left[\begin{array}{c} M^2 \\ (M \quad \mu) \end{array} \right] \quad \text{situación ideal}$$

(se licencia en la línea 2)

$$(\mu \quad \mu) \quad \mu$$

me n sa



TROQUEO SIMPLE

$$\left[\begin{array}{c} M \\ (\mu) (\mu) \end{array} \right] \quad \text{por defecto}$$

(licenciado en línea 1)

(ro) (sa) dos paréntesis (dos núcleos)



TROQUEO REDUCIDO

$$\left[\begin{array}{c} M \\ (\mu \quad \mu) \end{array} \right] \quad \text{más simple aún:}$$

v (i r) un solo paréntesis

Pensándolo bien, en *mensa* hay dos grupos de paréntesis: uno es el paréntesis interno de sílaba (como el de *vir*), que es el que tenemos en *m(en)*; y el otro es el paréntesis que delimita sílabas, como lo hacía en *rosa = (ro) (sa)*, que es el que nos da *(men) (sa)*. Así que hay dos maneras de formar un troqueo sencillo (es decir, con una sola proyección). El troqueo simple de *rosa* (preferible) tiene dos núcleos,

con dos paréntesis. Y el troqueo reducido de *vir* tiene un solo núcleo y un solo paréntesis. El troqueo compuesto es una combinación, en dos niveles sucesivos, de esas dos posibilidades. Tiene dos proyecciones, pues.

Cuando tenemos una palabra bisílaba que tiene material suficiente para formar el troqueo compuesto (*mensa*), estamos ante el ejemplo canónico. Pero, ¿qué hacemos con las palabras de más de dos sílabas? Empezamos por analizar las palabras de tres sílabas. Es de sobra conocido que hay dos posibilidades aquí. Si σ_2 es pesada, recibe el acento, pero si es ligera, el acento se traslada a σ_3 .⁹ Cuando σ_2 es pesada, simplemente tiene la estructura suficiente para formar un troqueo compuesto. Por ejemplo, *ami:cus* o *argentum*:

(13)

$$\begin{array}{cc} \left[\begin{array}{c} M^2 \\ (M \quad \mu) \end{array} \right] & \left[\begin{array}{c} M^2 \\ (M \quad \mu) \end{array} \right] \\ (\mu \mu) \mu & (\mu \mu) \mu \\ a \text{ mi } i \quad \text{cus} & ar \text{ ge } n \quad \text{tum} \end{array}$$

Ahora bien, cuando la σ_2 es ligera, no puede formar una supermora, y por tanto se apoya en σ_3 , que le aporta la mora necesaria. Por ejemplo, en *dominus* ‘dueño’, ‘señor’:

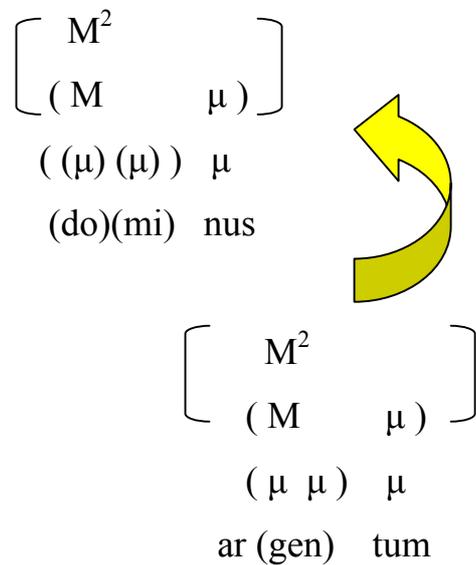
(14)

$$\begin{array}{c} \left[\begin{array}{c} M^2 \\ (M \quad \mu) \end{array} \right] \\ ((\mu) (\mu)) \mu \\ do \text{ mi } \quad \text{nus} \end{array}$$

Obsérvese cómo estamos siguiendo el camino inverso al que hicimos en sentido descendente. Al pasar de *mensa* a *rosa* prescindíamos de un nivel de paréntesis (el de la supermora), y al descender un nivel más hasta *vir* aún prescindíamos de un segundo nivel de paréntesis, el de los nodos silábicos. Pasábamos de los dos paréntesis de *(ro)(sa)* al único paréntesis de *(vir)*.

Ahora el camino es al contrario. Cuando no tenemos material suficiente, pero estamos operando en sentido ascendente en la estructura, en un nivel mayor de complejidad, pasamos de un solo paréntesis en *a(mi:)cus* o *ar(gen)tum* a dos paréntesis a la hora de formar la supermora en *((do)(mi))nus* y en *do((mes)(ti))cus* ‘casero’:

(15)



O sea que hemos formado un troqueo compuesto, pero añadiendo un nuevo nivel de complejidad estructural. Hay que entender que ante el dilema de una σ_2 ligera podríamos habernos conformado con un troqueo simple *do[mínus]* como el de *rosa*, que salvaba el problema.

Pero el troqueo ideal para un nombre es el compuesto, y por eso se prefiere aumentar la complejidad estructural, incorporando una nueva sílaba al recuento. Desde este punto de vista, podríamos estar hablando de una **captura de mora**.¹⁰ Se recluta una nueva mora en el núcleo previo, y así se puede formar una supermora que sirva de cabeza al troqueo compuesto. Como la propia supermora es de cabeza a la izquierda, será σ_3 quien porte el acento (*dóminus*) en la estructura (X (xx)).

Por cierto que al ser σ_3 un coadyuvante, no tiene trascendencia que su estructura interna sea “pesada” o “ligera”. Porque nos estamos moviendo en un nivel más alto, y por entonces ya ha pasado el momento de los paréntesis internos (de sílaba), que sí que importaban en la zona caliente de la σ_2 (*ami:cus*, *argentum*, *mensa*). Pero en *dominus* no importa, por estar un nivel más arriba:

(16)

$$\begin{array}{cc}
 \left[\begin{array}{cc} M^2 & \\ (M & \mu) \end{array} \right] & \left[\begin{array}{cc} M^2 & \\ (M & \mu) \end{array} \right] \\
 (\mu \mu) \mu & ((\mu)(\mu)) \mu \\
 a(mi i) cus & (do)(mi) nus
 \end{array}$$

Al tener dos paréntesis en *((do)(mi))nus*, la estructura interna ya no cuenta, frente al único paréntesis de *a(mi:)cus*, donde aún podemos abrir la transparencia y desglosarlo en *a(m(ii))cus* o en *ar(g(en))tum*. Pero en *dominus* ha pasado el momento. Por eso da igual la σ_3 “pesada” de *domesticus* que la “ligera” de *dominus*.

También da igual el “peso” (que en rigor no es tal peso, sino la cantidad) de la σ_1 a la hora de servir como complemento o ‘resto’ del troqueo compuesto. En tal situación, sólo se puede computar una mora (también le ha pasado el momento de acceder a su estructura).¹¹ Por eso no cambia nada entre *dómina*, *dóminus*, *dóminans*, porque la σ_1 siempre aporta una sola mora. [*domina* significa ‘dueña’ o ‘señora’, y *dominans* es ‘dominante’.]

La idea de capturar una mora parece tener otro ámbito de operaciones en la lengua latina. Todo sea por **reforzar** a σ_2 , la **sílaba pivote** del troqueo (tanto silábico como moraico). Igual que se captura un núcleo a la izquierda de σ_2 (en el nivel más alto), podemos también detenernos un nivel antes y capturar otra mora allí. Esto es lo que se hace cuando la σ_1 tiene un **inicio ramificado**, que al ser complejo permite ceder uno de sus segmentos. El caso más claro parece ser el de *cólubra* → *colúbra* ‘culebra’ que se hizo en el paso del latín clásico al latín vulgar, como recogen Szigetvári & Scheer (2005) (citando a Fouché 1966-73 I: 151s. y Pope 1934: 100).¹² Podemos postular que hubo una captura de mora de (co)(lu)(bra) a (co)(lub)(ra) en el plano métrico. Al tener una secuencia o cúmulo de dos consonantes, podemos capturar una de ellas (pero la otra se queda en el inicio de la σ_1). Y el fenómeno parece extenderse a los inicios ramificados con semiconsonante. Lo vemos en *memoria* ‘memoria’, que debe de ser *me(mor)ja*,¹³ y en los ejemplos de Steriade que cita Kenstowicz (1993: 574), *Músa* → *Musáque* ‘y la Musa’ y *límina* → *limináque* ‘y umbrales’, que nosotros interpretamos como *mu(sak)we* y *limi(nak)we* en el plano métrico.

Hasta aquí tenemos todo lo relativo al ámbito propio del acento, que según hemos planteado es un troqueo. En este enfoque, el acento no se asigna a una sílaba (directamente) como se había propuesto, sino a todo un pie. Por eso la σ_1 no es nunca extramétrica. Y gracias a que las moras quedan por encima del nodo σ , como interfaz entre sílabas y plano métrico, el peso puede trasladarse directamente y fielmente de un plano a otro. En cambio, cuando se ponían las moras por debajo de la sílaba, el plano métrico no las “veía”, y lo que veía eran nodos σ .

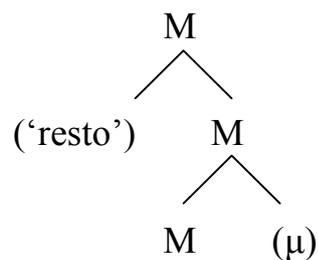
Una vez que el troqueo está perfectamente formado y tenemos su estructura proyectada, estaremos en condiciones de recibir el acento de palabra. Esto implica subir un nivel más arriba, al **ámbito de la palabra**. Aquí se mantiene la continuidad (es el troqueo el que proyecta al acento), y también debe mantenerse la distinción (la asimetría) entre cabeza y complemento. El troqueo es la cabeza de la palabra, y lo demás queda como ‘resto’. Como hemos dicho que el acento recae sobre un troqueo final, nos encontramos con que el ‘resto’ queda ahora a la izquierda de la cabeza.

Si no hay sílabas a la izquierda del acento, no hay ‘resto’. Se proyecta de forma **vacua** la cabeza del troqueo a cabeza de la palabra, y ya está. Pero si hay sílabas a la izquierda del acento, hay que delimitar dos ámbitos: cabeza (el troqueo mismo) y ‘resto’. Pudiera pensarse que ahora estamos ante una estructura de cabeza a la derecha, en franca contradicción con la cabeza izquierda que llevamos proyectando desde el nivel silábico (el propio núcleo silábico también es así). ¿De repente hay que cambiar el sentido de la estructura?

No lo tomamos así. Lo que pensamos es que el ‘resto’ de palabra no es un complemento, sino un **especificador** del troqueo¹⁴. Ya hay un complemento dentro del troqueo principal, y lo que queda fuera de él es su especificador. Así tenemos la proyección completa. La supermora M (también M² en el troqueo compuesto) es la cabeza métrica, y por eso es la que proyecta el acento de palabra. El material que tiene a su derecha es el complemento, según hemos dicho ya. Este complemento se proyecta dentro del propio troqueo, y está restringido a una sola mora. Pero en cambio, el especificador tiene un número muy libre de sílabas o moras, que va desde cero (cuando sólo existe el troqueo principal) a cualquier otra cifra (pensemos en *supercalifragilisticoespialidóso*).

Por ende, la palabra tiene una estructura de tipo **X-barra** con la supermora como cabeza, el complemento como tal (limitado a una mora), y el ‘resto’ como especificador:

(17)



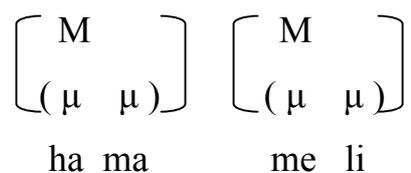
Esta estructura con formato X-barra es análoga a la que venimos utilizando para la sílaba. Al igual que el inicio no proyecta peso métrico, el especificador del troqueo no proyecta acento, y por eso queda relegado.

Vamos a ver cómo se asigna un acento en palabras de más de dos sílabas. El caso más sencillo es el de una palabra trisílaba pero con acento paroxítono, como eran *amí:cus* o *argéntum*. La sílaba sobrante a la izquierda del acento es su especificador. Lo mismo ocurriría en *canícula* ‘perrita’. Nos encontramos con que al llegar a este nivel, la proyección de moras sensibles ha pasado ya, y por lo tanto, las sílabas del especificador son por definición todas ligeras (tienen proyección semitransparente a los efectos del peso).

Nos fijamos ahora en una palabra más larga, el término neolatino *hamamelidanthemum* (que es una flor, ya extinguida). El troqueo se formó con *-danthemum*, y ahora el ‘resto’ es *hamameli-*.¹⁵ Este resto no se había computado hasta el momento, porque la zona caliente estaba en el troqueo compuesto. Solamente ahora que el troqueo principal está formado, y así la palabra tiene cabeza, nos ocupamos del especificador. Nos hemos abstenido de asignarle estructura hasta ahora (sólo como ‘resto’), en una suerte de *procrastination*.

Para dar estructura a *hamameli-*, lo hacemos de la forma más sencilla posible. Es lo más ‘barato’. Empezamos de izquierda a derecha, ya que la cabeza sigue estando, como siempre, a la izquierda. De este modo obtenemos sendos troqueos simples:

(18)



Después proyectamos un nivel más arriba, siempre con la cabeza a la izquierda:

(19)

$$\begin{array}{c} M^2 \\ (M \quad M) \\ \text{hama} \quad \text{meli} \end{array}$$

y así es como resulta la distribución de acentos, secundario y terciario. Recordemos una vez más que no afecta la “pesadez” de las sílabas, porque estamos en el especificador, y sólo rige la semitransparencia¹⁶ (la hora de la sensibilidad al peso ya pasó, cuando estábamos operando en la zona caliente del troqueo, en su cabeza)¹⁷.

Pensamos que este mismo patrón se aplica en inglés, y por eso en *Winnipesaukee* tenemos un troqueo final (-*saukee*) y un especificador a su izquierda, *Winnipe-*, que se lee de izquierda a derecha para asignar troqueos. Para que no se quede descabalada (sin licenciar métricamente) la sílaba final de *Winnipe-*, iteramos la proyección del troqueo simple (*Winni-*) para que la sílaba suelta *-pe-* se **absorba** dentro de un troqueo compuesto:

(20)

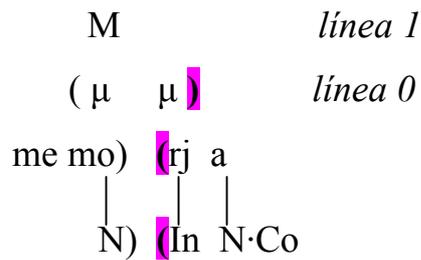
$$\begin{array}{c} \left[\begin{array}{c} M^2 \\ (M \quad \mu) \end{array} \right] \\ (\mu \mu) \quad \mu \\ \text{Wi nni} \quad \text{pe} \end{array}$$

De modo que hemos formado un troqueo compuesto en lugar del troqueo simple que se formaría por defecto al leer (*parse*) el especificador. Lo que no se forma, desde luego, es un **pie degenerado** (Hayes 1995), y tampoco hay que borrar ni reasignar nada, como se decía en Kiparsky (1979) (citado en Kenstowicz 1993: 576)¹⁸.

Como argumento final, planteamos que la captura de mora puede explicarse en consonancia con la estructura. Las dos capturas (la infrasilábica de *memoria* y la trocaica de *dominus*) pueden subsumirse en una sola por la posición de especificador que ocupa el material capturado. La /r/ de *me(mor)ja* se captura desde la posición de especificador silábico (el inicio, que además debe ser ramificado en esta circunstancia). Y la sílaba capturada en *(do)(mi)nus* también se captura desde la posición de especificador, sólo que ahora está en un nivel más alto, como especificador del troqueo. Por eso su estructura interna (silábica) ya no es accesible a la proyección. Un dato de interés es que para capturar material métrico no se “roba” de ninguna parte. Esto es, cuando un material pasa a recibir peso métrico, es porque antes no lo estaba recibiendo, no estaba entrando en el cómputo de su lugar de origen. Sólo así es posible la captura, tanto en el nivel infrasilábico (el inicio) como en el nivel trocaico (el ‘resto’).

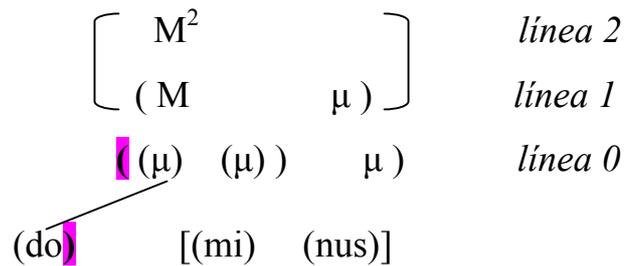
Al comparar las dos capturas de mora gráficamente, comprobaremos que la clave está en la falta de alineación entre el paréntesis del nivel inferior y el de nivel superior. Vamos a marcar en color esa falta de alineación, para que se aprecie más fácilmente. En la **captura infrasilábica** (*me(mo)ria* → *me(mor)ja*) la falta de alineación se da entre el paréntesis del nivel inferior (In (rja y el de nivel superior μ):

(21)



Por su parte, en $(do)(mi)nus \rightarrow ((do)(mi))nus$ hay una proyección fundamental del troqueo (cabeza + complemento), que marcamos entre corchetes [(mi) (nus)], y lo que está por fuera es el especificador:

(22)



Así ocurre la **captura trocaica**. De nuevo resaltamos la falta de alineación (marcada en color) entre el rótulo silábico de do-), que está fuera del corchete que marca el troqueo, y la estructura de moras donde μ ha capturado a la sílaba dentro de su ámbito.

Es más, nos planteamos que la captura trocaica pudiera hacerse ‘en segunda intención’, a saber, primero formando un troqueo simple y después implementándolo con la captura de la sílaba σ_3 . Este sería el proceso:

(23)

$$\begin{array}{r} \left[\begin{array}{c} M^2 \\ (\mu \ M) \ \mu \end{array} \right] \quad \textit{l\i{nea 2} \\ (\mu) \ [(\mu) \ (\mu)] \quad \textit{l\i{nea 1} \\ (\textit{do}) \ (\textit{mi}) \ (\textit{nus}) \quad \textit{l\i{nea 0} \end{array}$$

Donde la mora de complemento del troqueo simple (la de *-nus*) se ha vuelto a proyectar un nivel mas arriba cuando se incorpora la mora capturada de *do-*.

En conclusi3n, hemos dado cuenta de la asignaci3n del acento en latın con un solo esquema trocaico, del tipo X (x x). Gracias a una serie de mecanismos, hemos sido capaces de abarcar una gran variedad de patrones que a simple vista podıan parecer distintos. Cuanto mejor est3 articulado dicho mecanismo, mas productivo podra resultar, y con s3lo usar reglas sencillas. Por ejemplo, hemos derivado el troqueo compuesto a partir de la iteraci3n de estructuras en un troqueo simple. Eso sı, hay que licenciar el pie como tal, y ası es como distinguimos una humilde supermora de un troqueo simple.

Nuestro modelo aporta la proyecci3n de moras por encima del nodo silabico, y en forma de interfaz entre sılabas y pies. El grado de transparencia es importante para sentar distinciones. Con 3l se recoge la sensibilidad al peso de manera directa. Tambi3n hacemos una proyecci3n integral, que no pasa directamente de las sılabas al acento, sino que tiene en cuenta todos los niveles de la estructura. Hemos propuesto que ası puede evitarse la extrametricidad de la sılaba final, en tanto que esta sea el complemento de un pie trocaico. Podemos

además dar cuenta del ‘resto’ que queda a la izquierda del troqueo si lo tomamos como especificador del mismo.

¹ Esta característica se proyecta hacia arriba en la estructura, de manera coherente y continuada. Porque como dicen Szigetvári & Scheer (2005) (y nosotros suscribimos), el acento es una propiedad nuclear.

² Hayes (1989: 360) dice que *CVC syllables in Latin behave as heavy for purposes of stress, metrics, and other phenomena*.

³ Representamos la supermora de segundo nivel como M^2 , para distinguirla.

⁴ Ponemos la palabra “peso” entre comillas (y “ligera” o “pesada”) porque en sentido estricto nos estamos refiriendo a la **cantidad**, esto es, al número de elementos de que consta la sílaba.

Pero el **peso** es la proyección al plano métrico, y nada más, de acuerdo con nuestros términos, y en el troqueo simple dicha proyección es semitransparente, por lo cual la hemos representado con una sola mora.

⁵ La denominación de ‘**troqueo reducido**’ viene de asumir que un troqueo canónico consta de dos sílabas.

⁶ Esta es una decisión propia del latín, puesto que en árabe cairota sí que existen sílabas superpesadas, según Hayes (1995: 67-71, 89-90).

⁷ Sucede que la proyección de la sílaba es semitransparente. Y de nuevo nos encontramos con que el acento va a ser una propiedad nuclear.

⁸ En el troqueo simple, la cabeza es monomoraica por definición, ya que la proyección es semitransparente siempre. En el troqueo compuesto, sin embargo, la proyección es transparente (se computa el interior de la estructura silábica), y además hay un requisito de pesadez para poder formar la supermora.

⁹ Véase Hayes (1995: 91-92), por ejemplo.

¹⁰ La ‘captura de mora’ es un concepto que propusimos en el capítulo 9 de esta tesis.

¹¹ La semitransparencia se mantiene en secuencias de más una palabra. Por eso el verso cuantitativo latino considera “ligeras” las σ_1 de *úter* *vostró:rum* y de *dédit* *dó:no:*, según recoge Hayes (1995: 120) a partir de Allen (1973: 182-83). De acuerdo con nuestra idea, no es necesario un *iambic shortening*, como sostiene Hayes.

¹² Hay más ejemplos de este tipo: *ténebras* → *tenébras* ‘tinieblas’, *tónitrum* → *tonítrum* ‘trueno’, *íntegru(m)* → *intégru(m)* ‘entero’. Szigetvári & Scheer (2005) dan una explicación alternativa.

¹³ El interés de este ejemplo radica en que la /o/ de *memoria* es corta, y por ello la sílaba no estaría en condiciones de recibir el acento. Pero el hecho es que lo hace (es *memória*, y no *mémoria*), por lo que hay que pensar en la captura de mora. Un interés añadido de este ejemplo es que *memoria* es monomorfémica, con lo cual resta validez a los análisis cíclicos.

¹⁴ Al englobar al complemento y al especificador en la noción de ‘resto’, los diferenciamos de la cabeza. La distinción es útil, porque solamente la cabeza puede ser sensible al peso, es decir, al

material infrasilábico. Sólo la cabeza admite transparencia total. En cambio, los ‘restos’, tanto si se trata del especificador como si se trata del complemento, son por definición semitransparentes. Sólo pueden aportar una mora, que procede del propio nodo silábico, y el material o la estructura que haya por debajo no se tienen en cuenta.

¹⁵ Morfológicamente no es así, porque la palabra se formó por composición de *hamamelid* + *anthemum*. Pero fonológicamente su estructura es la que hemos analizado en el texto.

¹⁶ Combinando estructura métrica y semitransparencia explicamos la estructura del verbo *guberná:bunt*, con el acento secundario en *gù-* y la siguiente sílaba *-ber-* sin “pesadez” sólo es ligera porque pertenece al ‘resto’ (frente al *iambic shortening* de Hayes (1995: 120).

Y más todavía. Da la impresión de que la derivación tiene **efecto descendente**, en una especie de **formación regresiva** (*back formation*). (Recordemos que ya propusimos la formación regresiva como mecanismo fonológico en el capítulo 11 de esta tesis).

En vista de que las sílabas átonas se computan como ligeras en el nivel más alto, el métrico, se tiende a darles también ‘ligereza’ en el nivel silábico. Por eso se acortan las vocales de *égo*: → *égo* ‘yo’ y de *mále*: → *mále* ‘mal’, y tanto más aún cuando la vocal tónica es corta. Advirtamos que la cantidad se reduce a costa de la longitud vocálica, y no de eliminar un segmento de coda. (Los ejemplos son de Allen 1973, citados en Hayes 1995: 120).

También se reduce en *àmi:kítiam* → *àmikítiam* ‘amistad’, que Hayes compara con *amí:kus* ‘amigo’ (1995: 120). No es necesario invocar efectos cíclicos ni choque de acentos. Con el modelo que hemos propuesto se cubren estos ejemplos sin problema. Y no hay nada de *iambic shortening*, sino proyección semitransparente, que en ciertos casos parece tener un eco descendente, como acabamos de proponer.

¹⁷ Por cierto que esta estructura métrica del especificador, con la cabeza en el extremo izquierdo, remeda la que posiblemente se usaba en latín arcaico (véase Hayes 1995: 180). Es posible que el latín arcaico no tuviera especificador, y por eso el acento estuviera al principio de la palabra.

¹⁸ Hablando de Kiparsky, hay que explicar la diferencia entre *cond[ɛ]nsation* y *comp[ə]nsation*, (ver Kenstowicz 1993: 576). En nuestra opinión, lo que está en juego no es la **base**, como había dicho Kiparsky, sino la **raíz**. Y más concretamente, lo que importa es cómo se interprete esa raíz, y sobre todo de la posibilidad de reconocerla (recordemos el capítulo 10 de esta tesis). vemos cómo *con(dens)ation* remite fácilmente a *dense*, y a *density*, por lo que es reconocible por sí sola, mientras que *com(pens)ation* no puede referirse a ningún *pense*, por lo que *compens(ate)* necesita tomarse como un todo.

Analizamos un tercero en discordia: *concentration*. Aquí de nuevo se puede aislar o reconocer la raíz *center*, y por eso se escucha la vocal plena (o casi) cuando Jack Nicholson dice *con]-[centration* (así, en dos partes) en la escena de la máquina de escribir (*typewriter tantrum*) de ‘El Resplandor’: *Whenever you come in here and interrupt me, you’re breaking my con-concentration. You’re distracting me.*

Así que *concentration* tiene una base (*concentrate*) como la de *compensation* (su base es *compensate*), que además difiere de la base de *condensate*, que era *condense*. Pero el patrón vocálico de *concentration* es como el de *condensation* y no como el de *compensation*, por la diferencia de las raíces que hemos señalado.

Es más, al reconocer la raíz se tiende a separarla del prefijo [*con-*], y aquí de nuevo se agrupan *condensation* y *concentration*, donde se reconocen bien tanto la raíz como el prefijo *con-*. Por su parte, en *compensation* se ha asimilado el lugar de articulación del prefijo al de la raíz, por lo que es más difícil de reconocerlo como prefijo (*com-*) y separarlo de la raíz. Así que *compensate* es más propenso a tomarse como un todo (sin prefijo), y por eso forma un troqueo más regular, sin que la raíz *pense* tenga acento propio. Pero la primera sílaba de *condensation* tiende a quedar como prefijo (*con-*), y por eso vemos un mayor desglose de los componentes, y el posible acento terciario que sugería el SPE para *-den-*. La segmentación morfológica y la estructura de pies van de la mano, pues.

Y además de todo, la distinta formación de pies trae a colación los cúmulos consonánticos del contexto izquierdo que vimos en el capítulo 9. Ya entonces vimos la menor propensión a formar CS en cúmulos *-nd-* (por ejemplo, *tendon* o *Clarendon*), por mucho que sean homorgánicos. Por eso, aunque sea natural pronunciar *conc[ə]ntrate* y *conc[ə]ntration* con schwa, cuando hacemos una pronunciación marcada y separamos el prefijo, diciendo es mucho más probable que se escuche una vocal plena. Sin embargo, el cúmulo *-mp-* es mejor contexto izquierdo para formar una CS.

En definitiva, preferimos el ‘reconocimiento de la raíz’ y las diferencias de segmentación frente a la ‘derivación de la base’ que proponía Kiparsky.



Bibliografía

- Aitchison, J. (1991, 2001): *Language Change: Progress or Decay?* Cambridge: Cambridge University Press.
- Algeo, J. (2010): *The Origins and Development of the English Language*. (Sixth edition). Boston: Wadsworth, Cengage Learning (International Edition).
- Anttila, A. (2007): "Variation and optionality", en De Lacy (ed.), pp. 519-36.
- Baker, M. (1988): *Incorporation*. Chicago: The University of Chicago Press.
- Bald, W.-D. (1990): "An example of phonological reduction in English", en Ramsaran (ed.).
- Bell, A.: "Syllabic consonants", en Greenberg (1978), pp. 153-202.
- Blaho, S. (2004): "Syllabic consonants in Strict CV". MA thesis, Pázmány Péter University at Piliscsaba.
- Blevins, J.: "The syllable in phonological theory", en Goldsmith (1995), pp. 206-244.
- Blevins, J. (2004): *Evolutionary Phonology. The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.
- Bloomfield, L. (1933, reedición de 1961): *Language*. New York: Holt, Rinehart and Winston.
- Bošković, Ž. & H. Lasnik (eds.) (2007): *Minimalist Syntax. The Essential Readings*. Oxford: Blackwell.
- Brockhaus, W. (1995): "Skeletal and suprasegmental structure within Government Phonology". En *Frontiers of Phonology*, edited by Jacques Durand & Francis Katamba, 180-221. London & New York: Longman.
- Bronstein, A. J. (1990): "The development of pronunciation in English dictionaries", en Ramsaran (ed.).
- Burton-Roberts, N., P. Carr & G. Doherty (eds.) (2000): *Phonological Knowledge. Conceptual and Empirical Issues*. Oxford: Oxford University Press.
- Calabrese, A.: "Markedness Theory versus Phonological Idiosyncrasies," en Raimy & Cairns (eds.) (2009).
- Charette, M. (1998): "Empty and pseudo empty categories." *SOAS Working Papers in Linguistics and Phonetics* 8: 167-176.
- Chomsky, N. & M. Halle (1968): *The Sound Pattern of English*. New York: Harper and Row.
- Chomsky, N. (1995): *The Minimalist Program*. Cambridge, MA: MIT Press. Versión española de 1999: *El Programa Minimalista*. Madrid: Alianza Editorial.
- Chomsky, N. (1995): "Categories and Transformations", en *The Minimalist Program* (Cambridge, MA: MIT Press), en Bošković & Lasnik.

- Chomsky, N. (2000): *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.
- Chomsky, N., A. Belletti & L. Rizzi (2002): "An interview on minimalism", en *On Nature and Language*.
- Clements, G. N. (1985): "The Geometry of Phonological Features", en Goldsmith (1999), pp. 201-223.
- Clements, G. N. (2009): "The Role of Features in Phonological Inventories." En Raimy & Cairns (eds.) (2009).
- Clements, G. N. & S. J. Keyser (1983): "From *CV Phonology: A Generative Theory of the Syllable*", en Goldsmith (1999), pp. 185-200.
- Clements, G. N. & E. Hume (1995): "The Internal Organization of Speech Sounds", en Goldsmith (1995), pp. 245-306.
- Cruz Cabanillas, I. de la (2001): "Lexicología y semántica del inglés moderno", en I. de la Cruz Cabanillas & F. J. Martín Arista (eds.).
- Cruz Cabanillas, I de la & F. J. Martín Arista (eds.) (2001): *Lingüística Histórica Inglesa*. Barcelona: Ariel.
- Cruttenden, A.(2001): *Gimson's Pronunciation of English*. London: Edward Arnold.
- De Jong, K., B-J. Lim, and K. Nagao (2001): "The interaction of syllabification and voicing perception in American English".
- De Lacy, P. (ed.) (2007): *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press.
- Dell, F. & M. El Medlaoui (1985): "Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber." *Journal of African Languages and Linguistics*, 7, 105-130.
- Epstein, S. D. & Seely, T. D. (eds.) (2003): *Derivation and Explanation in the Minimalist Program*. Oxford: Blackwell.
- Ewen, C. J.: "Dependency relations in phonology", en Goldsmith (1995), pp. 570-585.
- Ewen, C. J. & H. van der Hulst (2000): *The Phonological Structure of Words. An Introduction*. Cambridge: Cambridge University Press.
- Fitch, W.T., M.D. Hauser & N. Chomsky (2005): "The evolution of the language faculty: Clarifications and implications". *Cognition* 97, 179-210.
- Fudge, E. C. (1969): "Syllables", en Goldsmith (1999), pp. 370-391.
- Giegerich, H. J. (1995): *English Phonology. An Introduction*. Cambridge: Cambridge University Press.
- Gimson, A. C. (1989): *An Introduction to the Pronunciation of English*. London: Edward Arnold.
- Goldsmith, J. A. (1989): *Autosegmental and Metrical Phonology*. Oxford: Blackwell.
- Goldsmith, J. A. (ed.) (1995): *The Handbook of Phonological Theory*. Oxford: Blackwell.
- Goldsmith J. A., (ed.) (1999): *Phonological Theory: The Essential Readings*. Oxford: Blackwell.
- Greenberg, J. (ed.) (1978): *Universals of Human Language. Volume 2: Phonology*. Stanford: Stanford University Press.
- Guerssel, M. (1986): "Glides in Berber and syllabicity". *Linguistic Inquiry* 17: 1-12.
- Haegeman, L. (1991, 1994): *Introduction to Government & Binding Theory*. Oxford: Blackwell.

- Hammond, M. (1999): *The Phonology of English. A Prosodic Optimality-Theoretic Approach*. Oxford: Oxford University Press.
- Hale, K. & S. J. Keyser (2000): "Conflation", en A. Bravo Martín, C. Luján Berenguel & I. Pérez Jiménez (eds.), *Cuadernos de Lingüística VII 2000, Documentos de Trabajo. Lingüística Teórica*, pp. 39-76. Madrid: Instituto Universitario Ortega y Gasset.
- Hale, M. & C. Reiss (2000): "Phonology as Cognition". En Burton-Roberts et al.
- Hall, T. A. (2007): "Segmental Features", en De Lacy (ed.), pp. 311-34.
- Halle, M. (2002): *From Memory to Speech and Back. Papers on Phonetics and Phonology 1954-2002*. Berlín: Mouton de Gruyter.
- Halle, M. & K. N. Stevens (1991): "Knowledge of Language and the Sounds of Speech". En Halle (2002), 176-195.
- Harris, J. (1994): *English Sound Structure*. Oxford: Blackwell.
- Harris, J. (2004): "Release the Captive Coda: the foot as a domain of phonetic interpretation". En J. Local, R. Ogden & R. Temple (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology 6*, 103-129. Cambridge: Cambridge University Press.
- Haspelmath, M. (2002): *Understanding Morphology*. London: Edward Arnold.
- Hauser, M. D., N. Chomsky & W. T. Fitch (2002): "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" *Science* vol. **298**: 1569-1579.
- Hayes, B. (1984): "The Phonology of Rhythm in English". *Linguistic Inquiry* **15** number 1: 33-74. Cambridge, Mass.: MIT Press.
- Hayes, B. (1989): "Compensatory Lengthening in Moraic Phonology". *Linguistic Inquiry* **20**. 253-306. En Goldsmith (1999: 351-369).
- Hayes, B. (1995): *Metrical Stress Theory: Principles and Case Studies*. The University of Chicago Press.
- Isac, D. & C. Reiss (2008): *I-Language. An Introduction to Linguistics as Cognitive Science*. Oxford: Oxford University Press.
- Jackendoff, R. (1993): *Patterns in the Mind. Language and Human Nature*. Harvester-Wheatsheaf.
- Jackendoff, R. (2002): *The Foundations of Language*. Oxford: OUP.
- Jones, C. (1989): *A History of English Phonology*. London: Longman.
- Jones, D., A. C. Gimson, & S. Ramsaran (1988): *Everyman's English Pronouncing Dictionary*. London: Dent & Sons.
- Kahn, D. (1976): *Syllable-based generalizations in English phonology*. Tesis doctoral, MIT. New York: Garland Press, 1980.
- Kaye, J. (1990): "'Coda' licensing". *Phonology* **7**: 301-330.
- Kaye, J., J. Lowenstamm, & J.-R. Vergnaud. (1990): "Constituent structure and government in phonology". *Phonology* **7**(2): 193-231.
- Kaye, J. (1997): "Why this Article is not about the Acquisition of Phonology". *SOAS Working Papers in Linguistics and Phonetics* **7**, 209-220. [Es una versión en "wild PDF", esto es, con posibles diferencias en paginación y texto respecto a la versión publicada].
- Kessler, B. (1998): "Ambisyllabicity in the Language of the Rigveda". En artsci.wustl.edu, last change, 2004-08-27.
- Kiparsky, P. (1982): "From Cyclic Phonology to Lexical Phonology", en J. A. Goldsmith (ed.) (1999): *Phonological Theory: The Essential Readings*, pp. 34-62.

- Kiparsky, P. (1995): "The Phonological Basis of Sound Change", en J. A. Goldsmith (ed.) *The Handbook of Phonological Theory*, pp. 640-670.
- Kreidler, C. (2004): *The Pronunciation of English. A Course Book*. Oxford: Blackwell.
- Ladefoged, P. (2001a): *A Course in Phonetics*. Boston: Heinle & Heinle.
- Ladefoged, P. (2001b): *Vowels and Consonants*. Malden, Massachusetts: Blackwell.
- Ladefoged, P. (2003): "Las razones profundas de los universales fonéticos", en R. Mairal & J. Gil.
- Ladefoged, P. & I. Maddieson (1996): *The Sounds of the World's Languages*. Oxford: Blackwell.
- Langacker, R. (1987): *Foundations of Cognitive Grammar*, I. Stanford, Stanford University Press.
- Lasnik, H. & J. Uriagereka, with C. Boeckx (2005). *A Course in Minimalist Syntax. Foundations and Prospects*. Oxford: Blackwell.
- Lavoie, L. (2002): "Subphonemic and Suballophonic Consonant Variation: The Role of the Phoneme Inventory", *ZAS Papers in Linguistics* **28**: 39-54.
- Lieber, R. (2010): *Introducing Morphology*. Cambridge: Cambridge University Press.
- Lightfoot, D. (1998): *The Development of Language*. Oxford: Blackwell.
- Mairal, R. & J. Gil (2003): *En torno a los universales lingüísticos*. Cambridge: Akal.
- Mora Bonilla, J.C. (2003): "The Formation of Syllabic Consonants and their Distribution in British English". *Atlantis* 25.2: 97-112.
- Moreno Cabrera, J. C. (1994): *Curso Universitario de Lingüística General*. Tomo II: *Semántica, pragmática, morfología y fonología*. Madrid: Síntesis.
- Muthmann, G. (2002): *Reverse English Dictionary*. Berlin: Mouton de Gruyter.
- Prince, A. (1983): "Relating to the Grid", *Linguistic Inquiry* 14, 19-100.
- Radford, A. (1988) (reimpresión de 1995): *Transformational Grammar. A First Course*. Cambridge: Cambridge University Press.
- Radford, A. (2004): *Minimalist Syntax*. Cambridge: Cambridge University Press.
- Ramsaran, S. (ed.) (1990): *Studies in the Pronunciation of English. A Commemorative Volume in Honour of A. C. Gimson*. London: Routledge.
- Raimy, E. & C. E. Cairns (eds.) (2009): *Contemporary Views on Architecture and Representations in Phonology*. Cambridge, Mass.: The MIT Press.
- Ramachandran, V. S. (2011): *The Tell-tale Brain. A Neuroscientist's Quest for What Makes Us Human*. New York: Norton.
- Real Academia Española (2011): *Nueva gramática de la lengua española. Fonética y fonología*. Barcelona: Espasa Libros S.L.U.
- Rennison, J. R. & F. Neubarth (2003): "An x-bar theory of Government Phonology", en *Living on the edge. 28 papers in honour of Jonathan Kaye*, ed. Stefan Ploch, 95-130. Berlin: Mouton.
- Rialland, A. (1986): "Schwa et syllabes en français", en L. Wetzels & E. Sezer (1986): *Studies in Compensatory Lengthening*. Berlin: de Gruyter.
- Ridouane, R. (2002): "Words without Vowels: Phonetic and Phonological Evidence from Tashlhiyt Berber". *ZAS Working Papers in Linguistics*, **28**, 2002, 93-110.
- Roach, P. (2000): *English Phonetics and Phonology*. Cambridge: Cambridge University Press.
- Roca, I. (1994): *Generative Phonology*. London: Routledge.

- Roca, I. (ed.) (1997): *Derivations and Constraints in Phonology*. Oxford: Blackwell.
- Roca, I. & W. Johnson (1999): *A Course in Phonology*. Oxford: Blackwell.
- Romano Mozo, M. (2001): "El léxico medieval inglés: características y evolución", en I. de la Cruz Cabanillas & F. J. Martín Arista (eds.).
- Rotshtein, P., R. N. A. Henson, A. Treves, J. Driver & R. J. Dolan (2005): "Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain", *Nature Neuroscience*. 8, 107-113.
- Sampedro, J. (2002): *Deconstruyendo a Darwin*. Barcelona: Crítica.
- Schane, S. A. (1995): "Diphthongization in Particle Phonology", en Goldsmith (1995), pp. 586-608.
- Scheer, T. (2003a): "Structure is a sovereign referee: the Coda Mirror." *CASTL kick-off conference*. Tromsø 2-4 October 2003.
- Scheer, T. (2003b): "Syllabic and trapped consonants in Slavic: different but still the same." *Formal Description of Slavic Languages 5*, Leipzig 26-28 November 2003.
- Scheer, T. (2004a): "A better solution for extrasyllabicity than extrasyllabicity." *GLOW 27*, Thessaloniki 19-21 April 2004.
- Scheer, T. (2004b): "Why parameters should be encoded in the software, rather than in the hardware." 12th Manchester Phonology Meeting 20-22 May.
- Scheer, T. (2004c): "How yers made Lightner, Gussman, Rubach, Spencer & Co. Invent CVCV". To appear in the proceedings of Generative Linguistics in Poland (GliP).
- Scheer, T. (2007a): "On the status of Word-Initial Clusters in Slavic (and Elsewhere)". En R. Crompton, M. Golezinska & U. Savchenko (eds.), *The Toronto Meeting (Michigan Slavic Materials 52)*, 346-364. Ann Arbor, MI: Michigan Slavic Publications.
- Scheer, T. (2007b): "How Semitic is Slavic? Initial clusters and syllabic consonants." Paper presented at the Formal Description of Semitic Languages 7. Leipzig.
- Scheer (2008): "Syllabic and Trapped Consonants in (Western) Slavic: the Same but yet Different". En Zybatow, G., L. Szucsich, U. Junghanns & R. Meyer (eds.): *Formal Description of Slavic Languages: The Fifth Conference, Leipzig 2003*, 149-167. Frankfurt am Main: Lang.
- Scheer, T. (2009): "Syllabic and trapped consonants in the light of branching onsets and licensing scales". En *Studies in Formal Slavic Phonology, Morphology, Syntax, Semantics and Information Structure*, edited by Gerhild Zybatow, Uwe Junghans, Denisa Lenertová & Petr Biskup, 411-426. Frankfurt am Main: Lang.
- Ségéral, P. & T. Scheer (1999): "The Coda Mirror" (version 3,1 sobre "Is the Coda Mirror a phonological object?"). Paper presented at the First Annual Meeting of the GDR 1954 "Phonologie" on Lenition and Fortition, Nice.
- Selkirk, E. O. (1982): "The syllable", en Goldsmith (1999), pp. 328-350.
- Spencer, A. (1995): *Phonology: Theory and Description*. Oxford: Blackwell.
- Spencer, A. (1996): *Morphological Theory*. Oxford: Blackwell.
- Szigetvári, P. & T. Scheer (2005): "Unified representations for the syllable and stress." *Phonology 22*: 37-75.
- The Advanced Learner's Dictionary of Current English*. London: Oxford University Press. Edición de 1960.

- Toft, Z. (2002): "From phonetics to phonology and back again: syllabic consonants in Southern British English." *ZAS Working Papers in Linguistics* **28**, 2002, 111-144.
- Webster's New Encyclopedic Dictionary* 1994 edition. Cologne: Könemann.
- Wells, J. C. (1990): "Syllabification and allophony", en Ramsaran (ed.).
- Wells, J. C. (1995): "New syllabic consonants in English". *Studies in General and English Phonetics: Essays in Honour of Professor J. D. O'Connor*. Ed. Jack Windsor Lewis. London: Routledge. 401-12.
- Yu, Alan C. L. (1999): "Aerodynamic constraints on sound change: the case of syllabic sibilants", *Proceedings of the International Congress of Phonetic Science*, San Francisco.